

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

**Defective images within this document are accurate representation of
The original documents submitted by the applicant.**

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SIANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 05:03:11 ; Search time 13643.2 Seconds
(without alignments)
862.151 Million cell updates/sec

Title: US-09-763-836-1

Perfect score: 713

Sequence: 1 gccagccctgatggggc.....atttgggtaaggtcatgat 713

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.om.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.sy.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htgo.hum.*
- 31: em.htgo.inv.*
- 32: em.htgo.rod.*
- 33: em.htg.hum.*
- 34: em.htg.inv.*
- 35: em.htg.rod.*
- 36: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Match	Length	DB	ID	Description
1	695.4	97.5	9616	14	AB049088	AB049088 Hepatitis
2	692.2	97.1	9427	14	HPCJRNA	D14484 Hepatitis C
3	692.2	97.1	9616	14	AF139594	AF139594 Hepatitis
4	692.2	97.1	9379	14	AF207766	AF207766 Hepatitis
5	691.6	96.9	9418	14	HCV132996	AJ132996 Hepatitis
6	690.6	96.9	1880	14	HFC5TRJ4	D08832 Hepatitis C
7	690.6	96.9	2540	6	E04260	E04260 cDNA encodi
8	690.6	96.9	2540	6	E04805	E04805 cDNA encodi
9	690.6	96.9	2540	6	E07391	E07391 cDNA encodi
10	690.6	96.9	9448	14	HPCJ483	D13558 Hepatitis C
11	690.6	96.9	9480	14	AB049094	AB049094 Hepatitis
12	690.6	96.9	9504	14	AB049099	AB049099 Hepatitis
13	690.6	96.9	9587	14	AF333324	AF333324 Hepatitis
14	690.6	96.9	9595	6	AR119832	AR119832 Sequence
15	690.6	96.9	9595	14	AF054247	AF054247 Hepatitis
16	690.6	96.9	9599	6	AR119833	AR119833 Sequence
17	689.4	96.7	9533	14	HCV132997	AJ132997 Hepatitis
18	689	96.6	780	6	AR095006	AR095006 Sequence
19	689	96.6	9456	14	HFCRNA	D10934 Hepatitis C
20	689	96.6	9460	14	HPCJ491	D10750 Hepatitis C
21	689	96.6	9595	14	AF054248	AF054248 Hepatitis
22	688.8	96.6	9379	14	AF165059	AF165059 Hepatitis
23	688.8	96.6	9379	14	AF165060	AF165060 Hepatitis
24	687.4	96.4	2033	6	E08374	E08374 N-terminal
25	687.4	96.4	2033	6	E08869	E08869 5'region of
26	687.4	96.4	4987	6	E08872	E08872 DNA encodin
27	687.4	96.4	9431	14	HPCPCPO	D45172 Hepatitis C
28	687.4	96.4	9431	14	HPCPP	D30613 Hepatitis C
29	687.4	96.4	9471	6	E08261	E08261 cDNA encodi
30	687.4	96.4	9471	6	E08457	E08457 cDNA encodi
31	687.4	96.4	9348	14	D89815	D89815 Hepatitis C
32	687.2	96.4	9379	14	AF165052	AF165052 Hepatitis
33	687.2	96.4	9400	14	HPCCGENOM	L02836 Hepatitis C
34	685.8	96.2	923	6	AR153761	AR153761 Sequence
35	685.8	96.2	9535	14	D85516	D85516 Hepatitis C
36	685.6	96.2	9377	14	AF207763	AF207763 Hepatitis
37	685.6	96.2	9379	14	AF165051	AF165051 Hepatitis
38	685.6	96.2	9410	14	HPCK1R2	D50481 Hepatitis C
39	684.8	96.0	9414	14	HCUKNCDS	M96362 Hepatitis C
40	684.8	96.0	9415	14	HCU16362	U16362 Hepatitis C
41	684.2	96.0	9431	14	HCU45476	U45476 Hepatitis C
42	684.2	96.0	9600	14	AF176573	AF176573 Hepatitis
43	684.2	96.0	9605	14	HCV238799	AJ238799 Hepatitis
44	684	95.9	9379	14	AF207758	AF207758 Hepatitis
45	684	95.9	9379	14	AF207767	AF207767 Hepatitis

ALIGNMENTS

RESULT : 1	AB049088	9616 bp	RNA	VRL	05-DEC-2000
LOCUS	Hepatitis C virus genomic RNA, complete genome, isolate: HCVT094.				
DEFINITION	AB049088				
ACCESSION	AB049088.1				
VERSION	GI:11559442				
KEYWORDS	Hepatitis C virus (isolate: HCVT094, specific_host: human) serum CDNA to genomic RNA.				
SOURCE	Hepatitis C virus				
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae; Hepacivirus.				
REFERENCE	1 (sites)				
AUTHORS	Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K., Hatahara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and Mishiro, S.				
TITLE	Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 9616)				

SOURCE Hepatitis C virus (strain:J33, specific_host:Homo sapiens) cDNA to genomic RNA.
ORGANISM Hepatitis C virus
VIRUSES: SRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 9427)
AUTHORS Cho, M.J.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-1991) to the DDBJ/EMBL/GenBank databases. M. J. Cho, National Cancer Center Research Institute, Virology Division; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104, Japan
(Tel:03-3542-2511(ex.4701), Fax:03-3542-2181)
REFERENCE 2 (bases 1 to 9427)
AUTHORS Cho, M.J.
TITLE Molecular cloning of Hepatitis C virus genome from a single Japanese patient
JOURNAL Unpublished (1993)
FEATURES Location/Qualifiers
1..9427
/organism="Hepatitis C virus"
/specific_host="Homo sapiens"
/strain="J33"
/db_xref="taxon:11103"
1..341
342..9374
/codon_start=1
/product="polyprotein"
/protein_id="BAA03375.1"
/db_xref="GI:1160328"
/translation="NSTNPKPQKTKRNTNRPODVPKPPGGQIVGGVILPPLRRPRL
GVATPKTSERQPRGRPIKARPEGRTWAOPGYDPLKNGSGMAGWALHSPRG
SRPSPGTPDRRSRLNGKVIDTLTCGADLMGYIPLVGLGAGRAALAHGRVLED
GVNATYATPCGFSIFILLALSLCTVPVSAYEVNRNVSAGYHVTNDCSNSSIVLPAAD
VLMHTPGVPCVRENSSRCWALPTLAARTNTVPTTIRHVDLLVGTAFACSAAY
VGDLCGSLVSQLTFSPRRHETVQDCNSIYPGHLSGRHAWDMNMSPTAALV
SOLLIPQVMDVAGAHWGLAGLAYSVMGNWAKVLI VMLFAGVDTGVTGAAV
ARTTGLHSFESAGSENLQINTNGSHINRTALNCDSLTGFLAALFTHKFNSS
RAESVLASCRIDEFDGNGPITYTERNSSDORPYCWHYPPKQGLIIPASEGVPYC
FPPSPVGTDFRFGVPTYSWENETDVLVLRNTPPOGNMFGCTMMNGTGTCTCGG
PCHTGGNNTLCTDTCFRKHPEATYKCGSGPMLTPRCMVDPYPLRHYPYCTVNF
TFKVRMYGVGVEHRLIAACNWTGRDNLDRSELSPLLSLSTEWQLIPCSFTTL
PALSTGLHLHNVDOVLYGIGSAVSIARWEYVLLFLLLADARCAVCLNWLML
TAQAAALENVLNVAASVAGAHGILSLFPCAANYIKGLRVPAAVAVLWQVWVPL
LLALPPRAYADREMAASCGGAVFMGLAFLTLSPHYKVFARLLWQLYLTRAEH
LHMVLPNVRGRDAIILLACAHPELIFIDITKULLALGLPLMLQAGITAVPYFVR
AQLGLRPLNVRKRVAGHYVQMAFKLAALTGYVYDHLTPLRDWAHAGLRDLAVE
PVFSDMETKITWGAADTAACGDIILGLPVSARRGRIILGLPADSLEGQWRLLAPIT
AYSOOTRGLGCIITSLTRDRKNOVEGVVYSTATQSLFATCINGVCTWYVHGAGSK
TLAGPKPITQWYTNVODLVGQWAPPGARSMTPCTCGSSDLVLYTRHADVTPVRRG
DGRGLSPRPVSYLKGSGGFLPCSGHANGIFRAAVCTRGVAKAVDEIPVESMETT
MRSPFTNNSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAAQGYKVLVLPVSAATL
GFGAYMSKAYGDPNVRGTVRTITGSPITYSTYCKFLADGCGSGGAYDIITCDRCHS
TDSTTILIGITVLDQAEAGARLVYATATPPGTVTPHPNIEEVLNTRGIPYK
ALPIEVITGGHRLIFCHSKKDELAALKSALGINAVAYRGLDVSVPASGDVVVVA
TDALMTGTGDRSDVIDNCNTQVDFSLDPTFTIETTVPQDASVSOBRGRTGRG
RAGIYRVTPCERPSGPDSSVLCYDSCGAWYELTFAETSVRLRAYLNPFLPVQC
DLHEFWESVFTGLTHDAFLQSOQAGNDFPILVAYQATVCARSOAPPSWDQWKK
LIRLKTLPHTPLLLYRIGAVONEVTLTHPTKIFIMACMSADLEVYVSTWVLGWA
ALAAVKTGSVIVGRIILSGRPVAPIDREVLVREFDEMECSHLPIEGMLLAE
QFKALGLLOWATKOAAEAPVETKQALEFWAKHWNFIISGIVLAGLSTLPGN
PAISIMAFTSITSLPTOSTLLENILGGVAAQLAPPASASAVGAGIAGNAVSI
GLGKVLVDLAGVAGVALVAFKMSGEVSPSTEDLVNLLPAILSPALVGVVCAA
LIRRHVPGEGAVQWNNRLIAFASRGNHVSPTVYVPSDAQRQVTOILSLTILQLK
RLUHWINDSTPCSGWLRDWMCTVLTDFKTWQLSKLLPLRPGLVPFPSCQSPYK
GVWRGEGTQMTTCCGCAQIAGHVKWMSMRI VGPRTCSNTHGTFPINAFTVTCSPGSK
APYASRLARVAAEVEYVTRVGFHYVTGVTVDNVCPCVOPAPFEETLDGVLRH
YAPACKPLLRDEVSQVGLNQYLSQLPCEPEPDAVLTSMLTDPSTHIAETAKRL
ARGPSPASSASQSLAFSLKATCTIHDSPDADLIEANLLRWEMGNINTRVSEEN
KVVIDLSPRIAEDEREVSPVAPILRRSRFPAMPPIWARPDPNPLLESWKQDPT
VPPVHGGCPPLPTKAAPIPPRPKRTIIVLSTESSALAEATKTFGGSGSAADSGT
ATATPDQTDSDGKESVSSMPLECPGDPDLSGWSSTVSEASDGVVCCMS
YTWTKLITPCAAKESKLPINPLSNLLRHNNVYATTSRASLQKQVTFDRLOVLD
DHYRDVLEMKAKASTVAKLLSVEACKLTPPHSKAFKFGYKADKDVLSLRVNHII

Source

5'UTR

CDS

TSVKKDLLEDTKTPTIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLGYRVCEKMALYD
VSTLTPQAVMSSGYQISPTGQRFVFLVTKWKKCPMGFSYDFRCSFTVSTVENDIRI
BESIYQCCDLAPAEKQVIRISLTERLYIGGPTNKGNGCYRRCSRAGVLTSCNTLL
TCYLKASAAACRAKQDCTMLVNGDGLVVICISAGTQEDAAASLRVTEAMTRYSAAPP
DPQPEYDLELITSCSSNVSAHDSGRVYVLTTRDPTPLAARAWEARHTPVNSWL
GNIIMVAPTLMARMLMTHFFSILLAOGLKALDQOIVGACYSIEPLDLPQI IERLH
GLSAPLSHSYSPGEINRVASCLRKLGVPPLRVHRARSVRKAKLSQSGRATCSKYL
FNWAVRTKLKLTPIPAASQDLSSWFAVAGYSGGDIYHLSLRAPRWFMLCLLLSVGV
GVYLLPNR"
9375..9427
BASE COUNT 1886 a 2821 c 2689 g 2031 t
ORIGIN
Query: Match
Best Local Similarity 97.1%; Score 692.2; DB 14; Length 9427;
Matches 700; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
1..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
2..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
3..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
4..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
5..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
6..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
7..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
8..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
9..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
10..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
11..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
12..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
13..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
14..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
15..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
16..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
17..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
18..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
19..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
20..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
21..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
22..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
23..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
24..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
25..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
26..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
27..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
28..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
29..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
30..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
31..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
32..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
33..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
34..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
35..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
36..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
37..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
38..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
39..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
40..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
41..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
42..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
43..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
44..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
45..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
46..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
47..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
48..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
49..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
50..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
51..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
52..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
53..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
54..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
55..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
56..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
57..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
58..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
59..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
60..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
61..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
62..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
63..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
64..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
65..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
66..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
67..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
68..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
69..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
70..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
71..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
72..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
73..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
74..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
75..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
76..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
77..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
78..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
79..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
80..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
81..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
82..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
83..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
84..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
85..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
86..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
87..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
88..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
89..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
90..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
91..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
92..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
93..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
94..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
95..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
96..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
97..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
98..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
99..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
100..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
101..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
102..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
103..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
104..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
105..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
106..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
107..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
108..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
109..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
110..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
111..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
112..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
113..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
114..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
115..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
116..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
117..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
118..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
119..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
120..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
121..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
122..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
123..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
124..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
125..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
126..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
127..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
128..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
129..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
130..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
131..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
132..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
133..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
134..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
135..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
136..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
137..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
138..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
139..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
140..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
141..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
142..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
143..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
144..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
145..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
146..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
147..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
148..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
149..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
150..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
151..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
152..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
153..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
154..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
155..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
156..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
157..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
158..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
159..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
160..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
161..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
162..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
163..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
164..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
165..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
166..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
167..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
168..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
169..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
170..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
171..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
172..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
173..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
174..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
175..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
176..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
177..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
178..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
179..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
180..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
181..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
182..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
183..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
184..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
185..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
186..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
187..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
188..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
189..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
190..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
191..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
192..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
193..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
194..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
195..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
196..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
197..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
198..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
199..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
200..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
201..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
202..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
203..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
204..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
205..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
206..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
207..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
208..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
209..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
210..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
211..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
212..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
213..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
214..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
215..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
216..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
217..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
218..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
219..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
220..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
221..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
222..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
223..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
224..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
225..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
226..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
227..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
228..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
229..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
230..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
231..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
232..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
233..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
234..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
235..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
236..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
237..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
238..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
239..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
240..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
241..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
242..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
243..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
244..1 gccagccctgatggggcgacactccaccatagatcaccctctgtgaggaactactg 60
|||||
245..1 gccagccctgatgg

KEYWORDS

SOURCE Hepatitis C virus.

ORGANISM

Hepatitis C virus.
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus.

REFERENCE

1 (bases 1 to 9616)

AUTHORS

Beard,M.R., Abell,G., Honda,M., Carroll,A., Gartland,M., Clarke,B.,

Suzuki,K., Lanford,R., Sangar,D.V. and Lemon,S.M.

TITLE

An infectious molecular clone of a Japanese genotype 1b hepatitis C

JOURNAL

Hepatology 30 (1), 316-324 (1999)

MEDLINE

99315771

PUBMED

10385673

REFERENCE

2 (bases 1 to 9616)

AUTHORS

Beard,M.R., Abell,G., Honda,M., Carroll,A., Gartland,M., Clarke,B.,

Suzuki,K., Lanford,R., Sangar,D.V. and Lemon,S.M.

TITLE

Direct Submission

JOURNAL

Submitted (30-MAR-1999) Microbiology and Immunology, The University

of Texas Medical Branch at Galveston, 301 University Blvd.,

Galveston, TX 77555-1019, USA

FEATURES

source

CDS

Location/Qualifiers
 1..9616
 /organism="Hepatitis C virus"
 /strain="HCV-N"
 /db_xref="taxon:11103"
 /note="infectious full-length clone
 genotype: 1b"
 342..9389
 /note="encodes structural and non-structural proteins"
 /codon_start=1
 /product="polyprotein"
 /protein_id="A044718.1"
 /db_xref="GI:5532422"
 /translation="MSTNPKPQRKTRNRPPQVDFPGGQIVGGVYLLPRGRPL
 SRAIRTSRSPQRQPIKARRPEGRWAQGYPPWLYNMGEGWAGWLVLRG
 SRPSWGTDPRRSRNLGKVIDTLTCLGLADLMGYIPLVGGPLGGAARALAGVRLSD
 GVNATNLPGCCSFIEALLISCLTPASAEHVRNAGSYHYHTNDCNSNIIPEAD
 LIMTGPCVFCRGNSRCRWALTPLAARNATLPTTIRHVDLLVGAALCSAMY
 VGDLCGSFVLSQFTSPRRHATLQDCNSIYPGHASHRMADMMNNSPTALVY
 SQLRIPDAVIDMAGHWGLAGLAYISMAGNWKVLIYMLLFGADHFTLTHGHA
 ARLTSGFAGLTTPGSQRIOLINTINGSWHNRNALNCNDSLOTGFLAALFYAIRFNS
 CDPKMSACRSIDKFDQGWPIYVAETKDPDORPCWYHPPQOCVPASOVCGPVY
 CTFPSVVTGDLRNPYSWGENPDLVLLANTRPQGNWRCCTWNSNIGTKTCG
 APCNIGCGNNTLTCTDCEFRKHEPATYKSCGSPWMLTFCRWVDYPIRLWHTPTCN
 FSIFKRMVYGGVHRUNAACNWTGRNCNDRSELSPLLSLTSTEWQVLPSCFTT
 LPALSTGLIHLQNIQVLYLGTSVAVFAIKWEYVLLFLLLADARVACLMMML
 LIAQEAALLENALNAASAGGAVFVGLALLTSLPYKVFARLLWLVLTTRAE
 LLILTLPPRAYAMDRNAASCGGAVFVGLALLTSLPYKVFARLLWLVLTTRAE
 HLHWVPLNAVGRGRDAIILLTCAVHPLELIDITKLIALILGLPLMWLAQGITRVPYV
 RAQGLIPLMLVRKAGGVVQAMFRLGALTGYVYNHLTPLRDWAHAGRLDLAV
 EPVFSDMETKIITWGDATACGDIILGLPVSARRREILLGPADSLVGGWRLLAPI
 TAYSQOTRGLGCIITSLTRGNQVVEGVVSTATQSFATCVNGVCWTVYHGAGS
 KTLAAGKPTAQMYTNDQVLVGPAPPGARSPTCTCGSSDLYLVTRHADYVTPRRR
 DNRGSLSPRPVSYLYKSSGGPLLCPSGHAVGVFRAAVCTRGVAKAVDPVPMET
 TMRSPTDNNSPAPVPTQVAHLHAPTSGKSTRVPAAYAAQGYKVLVNPVSAT
 LGFGAMSKAHGCTPNRTGRTITTCAPITYSTYKFLADGCGSGAYDILMCDECH
 STDSITLLGLTVDQETAGRLVALATAPPGSVTVPHNIEEVALSNTEIPFGY
 KATPIEAIGKRLLIFCHSKKCDLAAKSLGLGINAVAYIRGLDVSIVPTSGDVVYV
 ATDALMTGYTGDSDVIDNCTVQVDFSLDPTFTTETTPQDVAIRSQRRTGR
 GRGCIYRFTVTPGRSPGMSFVLCYCYDAGCAWELTPAETSVRLRAYRNTPLGVC
 ODHLERFESVFTGITHDAFLSQTQAGDNFYLVAOATVCARAQAPPSNDQWK
 CLIRLAKTLHGPTPLLYRLGAVQNEVTLTHPTKYIMACSADLEVYTWLVVGLV
 AALAAVCLITGVSVVVIGRIILSRGPVAVPREVLYREFDEMECSALHPYIEQMOLA
 EQFKALMLQLOATQAAAPVVEKSKRALETFWAKHKNFISGTQYLAGLSTLPG
 NPAISALMTASITSLPTNTLLNFGLVAAOLAPPASAASAFVAGIAGAAGVS
 IGLKVLVDILAGVAGVACALVAFKVMGSEAPSAEDLIALLPLSPGLVIVVCA
 AILLRHVHDECGAVQWNRLLAFACRGNHNSPHYYPDESAAARVQVLSLITQLL
 KRLHOWINEDCTPSCSWLURDWMVCTVLSDFKTLWQSKLLPLRPGLVPLFSCQY
 KGVWRGDMHTTCCPGAGTAGHVNGSMRIIGPKTCNTHWGTFTINAYTGTPTPS
 PAPNYSKALNRVAEEYEVTRVGDHYHTIITDNVPCQVPAPEFFTEVDGVRLLH
 PAPVCKPLRVAEEYEVTRVGDHYHTIITDNVPCQVPAPEFFTEVDGVRLLH
 LARGSPSLASSASOLAPSLRATCTTHSSYNLSDPDVLLIANLLWRCEMGNTR
 VESENKVVVLDSEPLRAEDENEISIAAEILLKSKFFPAIPIWAPDPNYPPLLESW

RESULT 4

AF207765

LOCUS

AF207766

9379 bp

RNA

VRL

27-APR-2000

KNPDYVPPVHVGCLPVPVKAIPPPIPRRRKRTVTLDTSTVSVLAELATKTFGSSELSA
 ADSCTATAPDQTDNGCKDSDBSCSMPLLEGEQDGPDLSDGWSVTSSEAGESVY
 CCSMSTWTGALITPCAAEESKLPINALSNLSLHHNVATTSSRAGLRQKKYTFDR
 LOVLDDHVRVWLKEMKASTVKAKLLISEACPLTPHSAKSKFGYCAKDVRLSSR
 AINHIRSVWEDLLEDTVPTDITVMAKNFVCQPEKGRKPARLIVFPDUGRVCEK
 MALYDVVSTILPQAVMGSSYGFQSPGQVBFVLKAWKSNKPMGFSYDTRCFDSTVTE
 NDRVESIYQCCDLAPARQAIKSLTERLYIGGLTNSKQSGCYRRCRASGVLTTS
 CGNFTLCYLKASACRAAKIQDCTMLVNGDVLVVICESAGTQEDAAISLRVTEAMTRY
 SAPPGDLPOPEYDELELTSSNSVVAHDASGKRVYVLTDRQPTPIARAAMETARHTP
 VNSWLGNIIMVAPTLMARMLITHFESILLAQEOLKALOCQIVGACYSLEPLDLPQI
 IERLHGLSARLSHSYSGEINRNASCLRKLGVPPLRVHRHARDVRAKLILSQGGAAT
 CGKYLFWNAWTKLKLTPIPAASOLDLSGMFVAGYSGGDIYHSLSRARPMFLCLLL
 LSVGVGIYLLPNR*

HASE COUNT 1891 a 2892 c 2734 g 2099 t

ORIGIN

Query Match 97.1%; Score 592.2; DB 14; Length 9616;
 Best Local Similarity 98.2%; Pred. No. 8.3e-167;
 Matches 700; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 gccagcccccgtgatggggcgacactccaccatagatcacctcccctgtgaggaactacg 60
 Dd 1 GCCAGCCCCGATTTGGGGCGACACTCCACCATAGATCCTCCCTGTGAGGAACACTG 60
 Qy 51 tcttcacgcagaaagcgtctagccatggcgttagtagtgatgtagtgcagcctccagac 120
 Dd 51 TCTTCACGCAGAAAGCGTCTAGCCATGGCGTTAGTATGATGATGATGATGATGATGAT 120
 Qy 121 ccccccctccgggagagccatagtgctgcggaaccggtgagtagacacgggaattgcag 180
 Dd 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCAG 180
 Qy 181 gacgacgggtcccttcttgatcaaacccgctcaatgctgagagattggcgctgcccc 240
 Dd 181 CACGACGGGTCTTCTTCTGATCAACCCCGCTCAATGCTCGAGATTTGGCGTGCCTCC 240
 Qy 241 gcgagactgctagcgcagtagtggtgggtgcgaaagccttggtgtagtgcctgtagg 300
 Dd 241 GCGAGACTGCTAGCGAGTAGTGTGGTTCGCGAAGGCCTTGTGTTACTGCTGATAGG 300
 Qy 301 gtgctgcgagtgccccggaggctctgtagacgtagcaccatgagacaaatccataac 360
 Dd 301 GTGCTTGCAGTGCCTCCGGGAGGTCTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAC 360
 Qy 351 ctcaagaaaaaacacgtaacaccaccgccccacagacgtaagttcccgggcg 420
 Dd 351 CTCAAAGAAAAACCAACGTAACACCAACCCCGCCCCACACGACGCTCAAGTTCCCGGCG 420
 Qy 421 gtgctcagatcgttggtgagtttacctgttgccgcgcagggggccccaggttggtg 480
 Dd 421 GTGCTCAGATCGTTGGTGGAGTTTACCTGTTGCCGCGCAGGGGCCCGCCAGTTGGGTGTC 480
 Qy 481 gcgcgactagaagacactccgagcggctcgcaacctctggaagggagacaaactatccca 540
 Dd 481 CGCGATCAGGAAGACTTCCGAGCGGTGCGAACCCCGTGGGAAGGCGACAGCTATFCCCA 540
 Qy 541 aggtcgc 600
 Dd 541 AGGCTCGCCGCGCGCGAGGCGGCGGCTGGGCTCAGCCCGGCTATCCTTGGCCCCCTCTATG 600
 Qy 601 gcaacgagggcagtggttgaggagtagtgctctctgtcgccgcgcgcgcgcgcgcgc 660
 Dd 601 GCAATGAGGCGATGGGTGGCAGAGTGGCTCTCTGTCACCCCGCGGCTCCCGGCGCTAGT 660
 Qy 651 ggggccccctcgacccccggcgtaggtcgtagtattgggttaaggtcatcgat 713
 Dd 651 GGGGCCCCACGACGACCCCCGCGCTAGGTGCGGCTAATTTGGTGAAGGTCTCATCAT 713

```

DEFINITION   Hepatitis C virus strain MD25 complete genome.
ACCESSION    AF207766.1  GI:7650249
VERSION      AF207766.1
KEYWORDS     Hepatitis C virus.
SOURCE       Hepatitis C virus.
ORGANISM     Hepatitis C virus.
              Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1 (bases 1 to 9379)
AUTHORS      Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and
              Sato, C.
TITLE        Characteristics of hepatitis C viral genome associated with disease
              progression
JOURNAL      unpublished
REFERENCE    2 (bases 1 to 9379)
AUTHORS      Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and
              Sato, C.
TITLE        Direct Submission
JOURNAL      Submitted (23-NOV-1999) Second Department of Internal Medicine,
              Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
              Tokyo 113-8519, Japan
FEATURES     Location/Qualifiers
              1..9379
               /organism="Hepatitis C virus"
               /strain="MD25"
               /db_xref="taxon:11103"
               /note="Isolated from a patient with liver cirrhosis and
               hepatocellular carcinoma (HCC-5)"
               330..9362
               /codon_start=1
               /product="polyprotein"
               /protein_id="AAF65956.1"
               /db_xref="GI:7650250"
               /translation="MSTNPKPQRTKRNRRPQDVKPPGGQIVGVYLLPRRGPRL
               SRVTRKTSRSPQRROPIPKARPEGRWQPGYWPPLYEGMEGWAQWLLSPRG
               GPRWGTRRRRNRLGKVIDTLTCGFDLMGTYPLVGPAGLGGAAALAHGVRLD
               GVNATLPGVCSFIFLLALLSCLTIPASAEVRNAGSVYHVDNCSNVSIEAAD
               MIMTPGCVPCVRNNSSRCWALTPTLAARNASVPTKTIIRHVDLLGAAAFASMY
               VGLDCGSVFLYSQLFTSPRHEVQDCNCISYIYGHLSGRHMDMMNNSPTAALV
               SOLLRIPQAMDVAGAHGVLGALYISVGNWAKVLIVMLFAGVDGSTRVSGGTA
               QRTTLVFAFFTSGPSQNIQVNTNGSHWIRNLTALNCNDSLOTGFAALFYTHFNDS
               GCPERMASCRSIDKEQGWGITVAGHSDLRPCYHYPQPGCIVPAQVCGPVYC
               FTFPSVVGTSRSGVPTYNGVNESDVLNLTNRPNQGNWFGCTWMSFTGTTCCG
               PPGNIGVGNNTLICPTDCEFRKHPDATYKCGSGPWLTPRCLVDYPERLWYPTVNE
               TILKRVYVGVGEVRLNAAWNTRGERCNLEDRSELSPLLSSTTEWQILPCSFETL
               PALSTGLIHQNVYQVLYIGISAVSVFVKWEYVLLFLFLADARVACLMMWLL
               IAQAPALEKLVILNAAASVAGAHGLSELVFCAAWYIKGLVPCAAAYLVGVNPLLL
               LLLALPPLAYANDREMAASCGGVFVGLVLLTSLPHYKVFRLARLWLIQYLITRAEL
               LQVWVPLNIRGGDAIILLTCAAHPELIFDISKILLAILGLPLVLOAGITRVPYEV
               AQLIRVCMVLRKVAGHYVOMALMKLAALTGTYYVDHLTPLRDWAHAGLRDLAVAVE
               VPISDMETKVTWCAPIAACVNIISGLPVSARKGKELLGPADSHESGRRLAPIT
               AYCOQTRGLLGCITISLTRDRKNQVEGEVQVSTATQSFATCVNGVCMVYHAGSK
               TLAGPKGTQMYINVDODLVGMVLPAGARSLLTPTCGSSDLYLIVTRHADYIPVRRG
               DSRGSLSPRVSVLKGSSGGLLPCPSGVVGIIFRAAVYTRGAKVAVDFVPEVSEMT
               TRSPVFTSDSPVAPVQTFVAHLHAPTGGVSKTPVPAAYAAQGVKVLVNPVVAATL
               FGFGAMSAHAGVDNIRGTITGAPITTYSTYKGLADGCGSGAYDIIMCDECHS
               TDSILGICVLDQATAGARLVVATATPGSVTVPHPNIERALSNTSEILPEYCK
               ALPIEAKGGRHLIFCHSKKCDLAAKLSGLNNAVYIRGLDVSIVPVSDEIVVVA
               TDALMTGTBGLSDVIDCNTQVDFSLDPTFTIETTPQDAVSRSRRGTRGR
               RRGYIRFTYPERSGMSVLCEDYDAGCAWELTPTAETSVRLRAYLNTPLGPVQC
               DHELFESVFTGLTHIDAFLSQTKAGDNLPLVAVOATVCARAOAPPSPDWQMKC
               LIRLPTLHGTPTLLYRLGAVONEVILTHPTTKYIMACMSADLELVSTVTVGSLVA
               ALAAVCLTGSVILVGRILLSRPVAVDREVLVREFDEMECSHSLPYIQGQOLAE
               QFKOKALMTOASTQKAAAPVSVESKRWLEAFWAKHWNFIISGIVLAGLSLTPGN
               PAIASLMAFTASITSPLTQSTLILNIIIGWVAQAOLAPPASAAPVAGAGAAGVSI
               GLGVLDILAGYAGVAGALVAFKMSGEMPESTEDVNLPLAIPSPALGVVGVCA
               ILRRHVGGEGAVQWMMRLIAFASRGNHVSPTHYVPESDAARVTOILSSLTITQLK
               RLHOWINEDCSTPGSGWLRDWMCTVLTDFKTLQSLKLLPLRGPVPLSCQGYK
               GYWRGDKLHHTICPGCAQITGHVKNRISIVGPRTCSNTHWGTIPVNAVITTPGCTPSP
               APNTSKALWAAVEYIEVTRVGFHYVMTDNTDNVCKPCQVPAPEFFTVDRVRLHR
               YAPACKPLREVEITFQVGLNQVLSQPCPEPDVAVLTSMLTDPSHITAEAKRL
               ARGSPPLSSASQLSAPSKATCTTRHSDPADLIEANLLWROEMGNITRVESEN
               KWVILDSPEPLRAEEDEREVSVAEEVLKRSRKEFPALPIWARPDPYNPPLLESWKDPDY

```

RESULT 5

HCV13296

LOCUS

HCV13296

9418 bp

RNA

VRL

04-MAY-1999

DEFINITION Hepatitis C virus, complete genome, isolate HCV-AD78.
ACCESSION AJ132996
VERSION AJ132996.1 GI:4753718
KEYWORDS complete genome; polyprotein.
SOURCE Hepatitis C virus.
ORGANISM Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 9418)
Rispeter.K.
AUTHORS Rispeter.K.
JOURNAL Thesis (1998) Universitaetsklinikum Essen, Institut fuer Virologie
REFERENCE 2 (bases 1 to 9418)
Rispeter.K.
AUTHORS Rispeter.K.
JOURNAL Direct Submission
TITLE Submitted (07-MAR-1999) Rispeter K., Universitaetsklinikum Essen,
JOURNAL fuer Virologie, Hufelandstrasse 55, 45122 Essen, GERMANY
INSTITUT Location/Qualifiers
FEATURES
source
1..9418
/organism="Hepatitis C virus"
/product="polyprotein"
/protein_id="CAB41950.1"
/db_xref="GI:4753718"
/translation="MSTNPKPORKTKRNTNRRPQDVKEFGGGQIVGGVYLLPRGRPL
CVARTKTSRSPRRRQKIPKARRPGRMAQGPMPLYNGLGKAGWGLSLSPRG
SRVSWPTDRRRNRLGKIDLTCTGDFADLMGYIPLVAGPLGGAARALAHGVRLSD
GVNATGPGCFSEFLLALLSCLTIPASAYEVNRVSGYHVHTNDCSSNIIYETAD
MIHTPGCVPCVSDNSRRCWALTPTLAARNGSVPTAIRRHVDLVLVGAACASAM
VGGCFSEFVLVSOLFSLPRRHETVOECNCSYIPGHVTHRMAMDMNKSPTLAVV
SOLLRIPOAVDMVAGAHGCVLAGLAYSMVGNKALLIYVLLFAGVGGGPTRTIGG
SQAASGLVSMFSGPSQKIQIINTNGSMHINRTALNDSLTGFLAALFAHFPN
SSCPERMASCRPIDFAQGWGPIITAEPSGLDQRPYCHWYAPQCGIPASVSCGPV
YCTPSPVVDTRDSGVPTIRYGTNETDYLNNTRPQGNWFGTWNITGFTKTC
GGPCNIGLGNNTLICPTDCKRHEATYKCGSGPMLTPRCIVDPYRLWHYPCVY
NFTIFKVMYVGCIEHRLSAACNWTGERCNEDRDRSELSPLLTSTWELIUCSFT
LLTAPSTGLIHONIVDQVLYGICGSLVSPAIKWEYILLPLLLADARVCACLNM
LLAQAEALNVLNLAASLAGAHLISFLFFCAAWYIKRLVPGAAVALYSWNL
LLLLALPRAYAMDRMAASGAGVGLVLLTSLPYKVLAKILWMLYFIRAE
AHLQVWIPPLNVRGGDAIILLTCAVHEPILFIDITKLLAILGLPLVLAQIIRAPYF
VRAQGLIRACMLRVAGGVYQMAFMKLAALTGYVYDHLTPLRDWAHAGRLDVA
VEPVFSQMETKIIITWAGDTAACDIIILGPLVSARRGKEILLGPADSLLEGWRLLAP
ITAYSDOTRGLGCIITSLTRDKNOVEGEVQVSTATOSFLATCVNGCWTVYHCG
SKTLAGKGPITOMYTNVDOLWGPAPSGARSILPTCTCGSSDLILVSHRADIVPVR
RDSRGLSLSPRVSYLKSSGGPILCPGSHVVGIFRAAVCTRGAKADVFPVESME
TITRSPVTDNNSPPAPVQTFQVHAHPTGSGKSTKFAAAYAAQYKVLVLPNSVAA
TLSFGAYMSKAHGVDPISIRTVRTITGAPITYSTYKFLADGSGGAYDIICDEC
HSTDSTSLIGTGLVDOATAGARLVLTATPPGSVTPVHPNIEVALSNTGEIPFY
GKAIPETIRGGRLHFLCHSKKCDLAAQLSGLINAVAYVRLDVSVPISGSDVVV
VATDALMTGDTGDSFDICNTCVTQVDESLDPTFIETTVPODAVSRQRRTG
RGRGRIYRTVTPGERSGDFDSVLECECDAGCAWELTPAETSVRKRLYTPGLPV
CQDLHFWESVFTGLTHIDAHFLSQKQADNPPYLVAYOATVCAARAAPPSSWDQW
KCLIRLPLHGHTPLLYRLGAVQNEVTLHPVTKYIMACMSADLEVVTSTVDOGM
LALAAAYCTGLTSVTVGRILLSGKPAIVPDGRVLYREFDEECASHLPYIEOGLML
AEQFKKALCLQATKQAEAAAPVVEKWRALFAWKHMMSFISGIOYLAGLSTLP
GNPAJASLMVAASVTSPLTTOHTLLFNILGGWVAQAQLAPPASAAPVAGTAGAAG
STGLKGLVDILAGYAGVAGALFAEKVMSGEMPTEDLVNLLPALISPCALVVGVC
AALLRRVDPGEGAVGMNRLALFASRGNHVSPTHYVPESDAARVTLISLTIQPL
LKLHQMINEEDSTPCGSLRDWDWICTVLTDFKTLQSKLLPLRPGVPFFSCRG
YKGVWRGDTGMOTTCGGAQITGHVKNRSMRIVGPRCTNSKTWHGTFPIINAYTGPCTP
SPAPYFRALRVAAEEYVEVTRVGDFFHVVTGMTDNKPCQVPAPEPFTVEDVGLR
HYAPACKPLLEVEVTFQVGLAQYLVGSQLPCEPEPGVTLTSMLTDPHSIATAETAKR
RLDRGSPSLASSASSQSLAPSKATCTTRHSDPDADLLEALLRWQEMGNITRVES
ENKVLVLDSPDLRAEEREISYPAEILRRSRKTPFRAMPIDNPYNPFLIESWKDP
DYVPVYVQGLPPAPKAPVPPRRRTVLTSTVSSALAEATLTFGSESSAVD
GTATAPGDSGDTGSDAESCSMPLEGPDPDLSDGWSWTVSEASDVEDVCCS
MSVYTGALITCAAESEKLPINALNSLLRHHNMVYATTSRASOROKKVFDRLOV
LDDHYWDVLKEMKAKASTVKARLLSVEEACKLTPPHSAKSKFCYGAADVRLNLSKAIN

BASE COUNT 1861 a 2843 c 2716 g 1998 t
ORIGIN
Query-Match 96.9%; Score 691; DB 14; Length 9418;
Best Local Similarity 98.6%; Pred. No. 1.7e-166;
Matches 697; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 7 cccctgatggggcgacactccaccatagatcaactccctctgaggaactactctctca 66
DB 6 CCCCTGTTGGGGCGGACACTCCACCATAGATCACTCCCTCTGAGGAACACTCTCTTCA 65
QY 67 cgcagaagcgtctagccatggcgttagtagtgatgctgacagcctccagagaccccc 126
DB 66 CCGAAGACGCTAGCCATGGCGTTAGTAGTGTGTCGACGCTCCAGAGACCCCCC 125
QY 127 tccgggagagcactatggtctctggaacccggtgagtacaccggaattgcccgcgcagac 186
DB 126 TCCGGGAGAGCCATAGTGGTCTCGCGAACCGGTGAGTACACCGGAATGCCAGAGAC 185
QY 187 cgggtctcttcttggatcaaaccccgctcaatgctggagattggcgctgcccgcgcgaga 246
DB 186 CGGTCTCTTTCTTGGATCAACCCGCTCAATGCTGTGAGATTTGGCGCTGCCCGCGAGA 245
QY 247 ctctaaccagtagttagtcgacgaagccttggtaactgctgatagggtcctt 306
DB 246 CTGCTACCCAGTAGTGTGGTTCGCGAAAGGCGCTTGTGTGTACTGCTGTATAGGTGCTT 305
QY 307 gcgagtgcccgggaggtctctgtagaccgtgacacatgagcacacaaatctcaaacctcaa 366
DB 306 GCGAGTGCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCTTAACCTCAA 365
QY 357 gaaaaacaaacgtataac 426
DB 356 GAAAAACAAACCTTAACACCAACCGCCGCCACAGGACGCTCAAGTTCCCGGGCGGTGCT 425
QY 427 agatcgtgtgtgagattactctgtccgcgcagggccccagagttgggtgtgcgcgcga 486
DB 426 AGATCGTGTGTGAGTTTACTGTGTCCGCGAGGGCCCCAGGTTGGGTGTGCGCGCGA 485
QY 487 ctagggaagactccgagcgtgctcaacctgtagaagcgacacacacacacacacacacacac 546
DB 486 CTAGGAAGACTCCGAGCGGTGCGCAACTCGTGGAGGCGACACCTATCCCCAAGCTC 545
QY 547 gccggccggagggcaggacactggcctcagccccgggtatcctctggccctctatggcaagc 606
DB 546 GCCGGCCCGAGGGTAGGGCTGGGCTCAGCCCGGGTACCCTTGCCCTCTATATGGCAACG 605
QY 607 agggcattgggtggcaggatggtcctctgcccccgcgctcccgccctagttggggcc 666
DB 606 AGGCGCTGGGGTGGGACAGTAGGCTCTGTCAACCCCGCGGCTCCCGGCTCTAGTTGGGGCC 665
QY 667 cttcgaccctcccgctgagtgctcgtaatttgggtgaaggtcgcgt 713
DB 666 CCACAGACCCCGCGTAGGTGCGGTAAATTTGGTAAAGGTATCGAT 712
RESULT 6
HPC5TRJ4 1880 bp RNA VRL 16-MAY-1998
LOCUS Hepatitis C virus gene for structural protein, partial cds,
DEFINITION isolate:HC-J4.
ACCESSION D00832
VERSION D00832.1 GI:221513

KEYWORDS	structural protein.
SOURCE	Hepatitis C virus (isolate:HC-J4) cDNA to genomic RNA.
ORGANISM	Hepatitis C virus
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE	1 (bases 1 to 1880)
AUTHORS	Okamoto,H., Okada,S., Sugiyama,Y., Yotsumoto,S., Tanaka,T., Yoshizawa,H., Tsuda,F., Miyakawa,Y. and Mayumi,M.
TITLE	The 5'-terminal sequence of the hepatitis C virus genome
JOURNAL	Jpn. J. Exp. Med. 60 (3), 167-177 (1990)
MEDLINE	91013116
REFERENCE	2 (sites)
AUTHORS	Okamoto,H., Okada,S., Sugiyama,Y., Kurai,K., Iizuka,H., Machida,A., Miyakawa,Y. and Mayumi,M.
TITLE	Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions
JOURNAL	J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)
MEDLINE	92044440
REFERENCE	3 (sites)
AUTHORS	Hotta,H., Doi,H., Hayashi,T., Purwanta,M., Soemarto,W., Mizokami,M., Oiba,K. and Homma,M.
TITLE	Analysis of the core and E1 envelope region sequences of a novel variant of hepatitis C virus obtained in Indonesia
JOURNAL	Arch. Virol. 136 (1-2), 53-62 (1994)
MEDLINE	94270990
REFERENCE	4 (sites)
AUTHORS	Mink,M.A., Benichou,S., Madaule,P., Tiollais,P., Prince,A.M. and Inchauspe,G.
TITLE	Characterization and mapping of a B-cell immunogenic domain in hepatitis C virus E2 glycoprotein using a yeast peptide library
JOURNAL	Virology 200 (1), 246-255 (1994)
MEDLINE	94174722
COMMENT	These data kindly submitted in computer readable form by: Hiroaki Okamoto Immunology Division, Jichi Medical School Kawachi-gun, Tochigi-ken 329-04 Japan Phone: 0285-44-2111 x3334 Fax: 0285-44-1557.

FEATURES	SOURCE
-----------------	---------------

CDS

```

/codon_start=1
/product="structural protein"
/protein_id="BAA00706.1"
/db_xref="GI:221514"
/translation="*NSTNPKPQKTKRNTNRNPDQVKFPGGGIIVGGVILLRPGRPL
GYRVRKTSERSQGWROPIPKARPEGRWAQPGYPWPLYGEGRAGWAGLLSPRG
SPRSPGTPRRSRNLKVIDTLCGFADLMGYIPLVGAPLGEARALAKHGVLELD
GYNATGNLPGCSFSFLALSLCLTIPASAEYVRNVSGIYHVNTDCSNSSIVYEARD
MIMHTPGVPCVFREDNSRCWALTPTLAARNASVPTTIRRHVDLLLVGAAPFCASMY
VDGCGSVLYSOLFSTSPRRHETVDCNCISYPGLSHGRMAWMDMMNSPTTIALY
SOLLTPQAVDMVAGAHGVLAGLAYIYSMVGNAKWLVIALLVFAGVDGTYTYSGGAA
SHITSTLASLFSFGASQRLQNTNGSIHNTALNCNDSLHTGFTLAAFLPTTYRRENS
GCPERMASCRPIDWFAQGGGPIITYTEPDSPDQRPYCHWYAPRPGCIVPASQVCGPYVC
PTSPD"

```

```

BASE COUNT      333 a      593 c      555 g      399 t
ORIGIN

Query Match      96.9%;      Score 690.6;      DB 14;      Length 1880;
Best Local Similarity 98.0%;      Pred. No. 2.8e-166;
Matches 699;      Conservative      0;      Mismatches 14;      Indels      0;      Gaps      0;

```

Qy 1 gccagcccccctgatgggggggacactccaccatagatcactccccctgtgaggaactactg 60
|||||
Db 1 gccagcccccCGATTgggggggacactccaccatagatcactccccctgtgaggaactactg 60
|||||

61	QY	tcttcacgcagaagcgtctagccattggtcgttagtatgtagtctgtagcagctccaggac	120
62	DB		
63	DB		
64	DB		
65	DB		
66	DB		
67	DB		
68	DB		
69	DB		
70	DB		
71	DB		
72	DB		
73	DB		
74	DB		
75	DB		
76	DB		
77	DB		
78	DB		
79	DB		
80	DB		
81	DB		
82	DB		
83	DB		
84	DB		
85	DB		
86	DB		
87	DB		
88	DB		
89	DB		
90	DB		
91	DB		
92	DB		
93	DB		
94	DB		
95	DB		
96	DB		
97	DB		
98	DB		
99	DB		
100	DB		
101	DB		
102	DB		
103	DB		
104	DB		
105	DB		
106	DB		
107	DB		
108	DB		
109	DB		
110	DB		
111	DB		
112	DB		
113	DB		
114	DB		
115	DB		
116	DB		
117	DB		
118	DB		
119	DB		
120	DB		
121	QY	ccccctcccggagagccatagtgctcgcggaaccggtgagtacacacgcggaattgcacg	180
122	DB		
123	DB		
124	DB		
125	DB		
126	DB		
127	DB		
128	DB		
129	DB		
130	DB		
131	DB		
132	DB		
133	DB		
134	DB		
135	DB		
136	DB		
137	DB		
138	DB		
139	DB		
140	DB		
141	DB		
142	DB		
143	DB		
144	DB		
145	DB		
146	DB		
147	DB		
148	DB		
149	DB		
150	DB		
151	DB		
152	DB		
153	DB		
154	DB		
155	DB		
156	DB		
157	DB		
158	DB		
159	DB		
160	DB		
161	DB		
162	DB		
163	DB		
164	DB		
165	DB		
166	DB		
167	DB		
168	DB		
169	DB		
170	DB		
171	DB		
172	DB		
173	DB		
174	DB		
175	DB		
176	DB		
177	DB		
178	DB		
179	DB		
180	DB		
181	DB		
182	DB		
183	DB		
184	DB		
185	DB		
186	DB		
187	DB		
188	DB		
189	DB		
190	DB		
191	DB		

RESULTS 7

E04260	2540 bp	RNA	PAT	29-SEP-1997
DEFINITION	CDNA encoding a part of type non-A non-B hepatitis virus.			
ACCESSION	E04260			
VERSION	E04260.1 GI:2172463			
KEYWORDS	JP 1933023200-A/2.			
SOURCE	unidentified.			
ORGANISM	unidentified.			
REFERENCE	1 (bases 1 to 2540)			
AUTHORS	Okamoto, H. and Nakamura, T.			
TITLE	HIGHLY SENSITIVE DETECTION METHOD OF NON-A NON-B TYPE HEPATITIS VIRUS; USING OLIGONUCLEOTIDE PRIMER AND OLIGONUCLEOTIDE PRIMER			
JOURNAL	Patient: JP 1933023200-A 2 02-FEB-1993:			

NAKAMURA TEISUO
PN JP 1993023200-A/2
COMPIED

PD	26-FEB-1993	
PF	26-APR-1991	JP 1991191376
PR	12-JUN-1990	JP 90P 153402
PI	KAMATO HIRAKI, NAKAMURA TETSUO	
PC	C12Q1/68,C12N15/51,C12Q1/70;	
CC	strandedness: Double;	
CC	topology: Linear;	
CC	hypothetical: No;	
CC	anti-sense: No;	
CC	*source: strain-HC-J4;	

```

FH      Key      Location/Qualifiers
FH      misc_feature 1..2540
FH      virus'
FH      Location/Qualifiers
FEATURES
  source      1..2540
              /organism="unidentified"
              /db_xref="taxon:32644"
BASE COUNT  470 a 776 c 741 g 553 t
ORIGIN
Query Match      96.9%; Score 690.6; DB 6; Length 2540;
Best Local Similarity 98.0%; Pred. No. 2.7e-166;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 1 gccagcccccgtatggggcgacactccaccatagatcacctcccctgtgagaaactactg 60
Db 1 gccagcccccgtatggggcgacactccaccatagatcacctcccctgtgagaaactactg 60
Qy 61 tcttcacgcagaaagcgtctagcattggttagttagtgcgtgcagcctccaggac 120
Db 61 tcttcacgcagaaagcgtctagcattggttagttagtgcgtgcagcctccaggac 120
Qy 121 ccccccctccggagagccatagtggtctgcggaaccggtagtacaccggaattgccag 180
Db 121 ccccccctccggagagccatagtggtctgcggaaccggtagtacaccggaattgccag 180
Qy 181 gacgacgggtcccttcttgatcaaccgcctcaatgctggaattggcggtgcccc 240
Db 181 gacgacgggtcccttcttgatcaaccgcctcaatgctggaattggcggtgcccc 240
Qy 241 gcgagactgtagcggagtagtgggtgcgaaagcctgtggtactgcctgatagg 300
Db 241 gcgagactgtagcggagtagtgggtgcgaaagcctgtggtactgcctgatagg 300
Qy 301 gtgcttgcgaatgccccggagggtctctgtagaccgtgcacacatgagacaatactaaac 360
Db 301 gtgcttgcgaatgccccggagggtctctgtagaccgtgcacacatgagacaatactaaac 360
Qy 361 ctcaaaagaaacaaacgtaacacacacccgcgcacagagcgtcaagtctcccgggc 420
Db 361 ctcaaaagaaacaaacgtaacacacacccgcgcacagagcgtcaagtctcccgggc 420
Qy 421 gtggtcagatcgttggtgagtttacctgttgcgcgcaggggcccaggttggtgtgc 480
Db 421 gtggtcagatcgttggtgagtttacctgttgcgcgcaggggcccaggttggtgtgc 480
Qy 481 gcgagactaggaagacttccgagcggctgcgaacctctgtggaagcgcacaacctatcccc 540
Db 481 gcgagactaggaagacttccgagcggctgcgaacctctgtggaagcgcacaacctatcccc 540
Qy 541 aggtctgcgcgcgcgcagggagccttggtctcagccgggtatccttggccccctctatg 600
Db 541 aggtctgcgcgcgcgcagggagccttggtctcagccgggtatccttggccccctctatg 600
Qy 601 gcaacgaggggcatgggtggcagatggctcctgtgcgcccgccggtcccgccgctagt 660
Db 601 gcaacgaggggcatgggtggcagatggctcctgtgcgcccgccggtcccgccgctagt 660
Qy 661 ggggccccacggacccccggcgtaggctgcgttaatttggttaaggctaatcagat 713
Db 661 ggggccccacggacccccggcgtaggctgcgttaatttggttaaggctaatcagat 713

```

RESULT 8

```

LOCUS      E04805      2540 bp      RNA      PAT      29-SEP-1997
DEFINITION cdna to 5'-terminal region of gRNA of Hepatitis nonA nonB virus.
ACCESSION  E04805
VERSION    E04805.1 GI:2173001
KEYWORDS   JP 1993091884-A/2.

```

```

SOURCE 1
ORGANISM
  unidentified.
  unclassified.
  1 (bases 1 to 2540)
  Okamoto, H. and Nakamura, T.
  DETECTION SYSTEM FOR NON-A NON-B HEPATITIS VIRUS RELATING ANTIGEN
  AND ANTIBODY, POLYNUCLEOTIDE AND POLYPEPTIDE
  Patent: JP 1993091884-A 2 16-APR-1993;
  NAKAMURA TETSUO
  OS Hepatitis nonA nonB virus
  PN JP 1993091884-A/2
  PD 16-APR-1993
  PF 10-APR-1991 JP 1991196175
  PR 12-JUN-1990 JP 90P 153401, 08-NOV-1990 JP 90P 304405 PI
  OKAMOTO HIROAKI, NAKAMURA TETSUO
  PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576,
  G01N33/577;
  CC strandedness: Single;
  CC topology: Linear;
  CC hypothetical: No;
  CC anti-sense: No;
  FH Key      Location/Qualifiers
  FT 5'UTR    1..341
  FT CDS      342..2540
              /note="protein coded by Hepatitis nonA nonB
              virus gene".
FEATURES
  source      1..2540
              /organism="unidentified"
              /db_xref="taxon:32644"
BASE COUNT  470 a 776 c 741 g 553 t
ORIGIN
Query Match      96.9%; Score 690.6; DB 6; Length 2540;
Best Local Similarity 98.0%; Pred. No. 2.7e-166;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 1 gccagcccccgtatggggcgacactccaccatagatcacctcccctgtgagaaactactg 60
Db 1 gccagcccccgtatggggcgacactccaccatagatcacctcccctgtgagaaactactg 60
Qy 61 tcttcacgcagaaagcgtctagcattggttagttagtgcgtgcagcctccaggac 120
Db 61 tcttcacgcagaaagcgtctagcattggttagttagtgcgtgcagcctccaggac 120
Qy 121 ccccccctccggagagccatagtggtctgcggaaccggtagtacaccggaattgccag 180
Db 121 ccccccctccggagagccatagtggtctgcggaaccggtagtacaccggaattgccag 180
Qy 181 gacgacgggtcccttcttgatcaaccgcctcaatgctggaattggcggtgcccc 240
Db 181 gacgacgggtcccttcttgatcaaccgcctcaatgctggaattggcggtgcccc 240
Qy 241 gcgagactgtagcggagtagtgggtgcgaaagcctgtggtactgcctgatagg 300
Db 241 gcgagactgtagcggagtagtgggtgcgaaagcctgtggtactgcctgatagg 300
Qy 301 gtgcttgcgaatgccccggagggtctctgtagaccgtgcacacatgagacaatactaaac 360
Db 301 gtgcttgcgaatgccccggagggtctctgtagaccgtgcacacatgagacaatactaaac 360
Qy 361 ctcaaaagaaacaaacgtaacacacacccgcgcacagagcgtcaagtctcccgggc 420
Db 361 ctcaaaagaaacaaacgtaacacacacccgcgcacagagcgtcaagtctcccgggc 420
Qy 421 gtggtcagatcgttggtgagtttacctgttgcgcgcaggggcccaggttggtgtgc 480
Db 421 gtggtcagatcgttggtgagtttacctgttgcgcgcaggggcccaggttggtgtgc 480
Qy 481 gcgagactaggaagacttccgagcggctgcgaacctctgtggaagcgcacaacctatcccc 540
Db 481 gcgagactaggaagacttccgagcggctgcgaacctctgtggaagcgcacaacctatcccc 540
Qy 541 aggtctgcgcgcgcgcagggagccttggtctcagccgggtatccttggccccctctatg 600
Db 541 aggtctgcgcgcgcgcagggagccttggtctcagccgggtatccttggccccctctatg 600
Qy 601 gcaacgaggggcatgggtggcagatggctcctgtgcgcccgccggtcccgccgctagt 660
Db 601 gcaacgaggggcatgggtggcagatggctcctgtgcgcccgccggtcccgccgctagt 660
Qy 661 ggggccccacggacccccggcgtaggctgcgttaatttggttaaggctaatcagat 713
Db 661 ggggccccacggacccccggcgtaggctgcgttaatttggttaaggctaatcagat 713

```


Hiroaki Okamoto
Immunology Division
Jichi Medical School
Kawachi-gun
Tochigi 329-04
Japan
Phone: 0285-44-2111 x3334
Fax: 0285-44-1557

FEATURES

Location/Qualifiers
1..9448

/organism="Hepatitis C virus"

/isolate="HC-J4/83"

/strain="HC-J4"

/db_xref="Caxon:11103"

342..9374

/codon_start=1

/product="polyprotein precursor"

/protein_id="BAA02756.1"

/db_xref="GI:221605"

/translation="MSTNPKPQRTKNNRRPQDVKFPQGGQIVGGVYLLPRGPRLL
GVRATRTSESRQGRWQIPKARPEGRWAGQPGWPLYGNEGGLWAGWLLSPRG
SRWSGPTDRRRNRGLKVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLED
GVNATGNLPGCSFIELLALLSCLTIPASAYEVNRYGVYHVTNDCSNSSIVYEAAD
MIMHTPCVPCVRDESSRCWVALTPTLAARNAVPVTTIRRHVDLLVGAARPCSAKY
VGLDGGSVFLVSOLFTEPRHETVQDCNCSYFGHLSGRHMAWMMNNSPTIALVY
SOLLRIQAVVDVAGAHGVLGATYISWGNNAKVLIVALLFAGVDETYTSGGAA
SHITSTLASFSQASQRIQLVNTNSWHINRTALNCNLSLHTGFLAALFYTHFNSS
CGPERMASTPITFAQWGPITYTEPDSQDPYCHWYAPRCGIPASQVCGPYVC
FTSPVVVGTDRSGVPTYSGWENETDMLNTRPPQGNWFGCTWNSVTFGTCTGG
PPCNIGVGNHTLCPTDCPRKHPEATYTKGSGPMLTLCPLVDYPIYRIWHYPCTEFP
SIPKRVYGVGEHRLNAACNMTGRCNLEDRDRSELSPLLSTTEWILLPCAFTTL
PALSTGLHLHQNIVDQYLYGVGSFVSAIKWEYILLFLLADARVACALMULL
LQAERALELVNVAASVAGHILSFLVFFCAAWYIKGRAPGAAYFYGVNPLLL
LLALLPRAYMDREMAASCGAVLGLVFLTLPSPYKVFELTRILMWOYFTRAEAH
QWVWVPLNRRGGDAIILLTCAVHPELIFDIKLLAILGLMLVLAQGITRIVPYFVR
AQGLIRACMLVRKVAAGHYVQMAFMKLCALTYVYNNHTPLRDWAHAGRLDAVVE
PVFSDMETKVTWAGTAACGDIILGLPVSARRGKEIFLGPADSLRQGNWRLAPIT
AYSOOTRGVLGCIITSLTGRDKQVEGEVQVSTATOSFLATCINGVCVYVYHAGAK
TLAGPKPIQMTNVDLDLVGQAPGASMTPTCTGSSDLXILTRHADVIVPVRRG
DSRGLSPRPVSLKSGSGPLLCPSGVVGVFRAAVCTRGAKAVDFTEPVESMET
MRSPFTDNSPPAPQTFQVAHLHAPTGSKSTKVPAAQYKVLNPSVAATL
GFGAMPKAGIDPNIRTVRTITTGPIITYSYGKFLADGGSGGAYDIIICDECHS
DTYITLIGIQLDIOAETAGARLVATATPGSVTVPHNIEIGLNNGETIPFYGK
ALPIEALIKGRHLIFCHSKKCDLAKLGLGNVAVYRGLDVSVIPIGDVAVVA
TDALMTGTGDSFVDCNTCVQTVDFSLDPTFTIETTVQDVAISRQRRTGKG
RSGLRIVTEPERSGMSDSCECYDAGCAWTELPATSVRLRAYLNTPLGPLVQ
DHLEFVESFTGLSHIDAHLFSQTKQAGDNFPLVAYQATVCAQAQPPSPWQMKC
LIRLPTLHGPTPLLYRLGAVQNEVTLTHPTIKYIMACMSADLEVVTSTWLVGGVLA
ALAACTLTQSVTVGRILISGKPAVVPDEVLYQEFDEMECASOLPYTEOGMLAE
QFKOKALLQATKQAEAAAPVSVESKRALETFWAKHWNFISSLOYLASLTPGN
PATASLMAFTASTISPLTQNTLLENLILGWVAAQLAPPASNAFVAGTAGRAVGI
GLKVLVDILGAGVAGVAGALVAFKVMGSEVPSTEDLNLPLPAILSPVGVVVCAA
ILRRHVGPGGAVQMMNRLIAFASRGNVSEPTEDLNLPLPAILSPVGVVVCAA
RLHOWINEDCTPCSGSLRDVMDICTVLDFKTLQSKLPLPLPGVPLFSQGRYK
GVNRGDIIMQTCPCGAQIAGHVKNKSMRIVGPTCSNTHWGTFFPNAVYTPCTSP
APNYSRALWMAAEYEVTVRGVDFHVVTGMTDNNVKPCQVPAPEFTTEVDGRVLRH
YAPACKLLREDVAFQVGLNOYLVGSQLPCEPEPDVTLTSLMIDPSHTAETAKRL
ARGSPPLSSASSQSLSAPSKATCTTHSDSPADLEIANLLWRQEMGNTIYVESEN
KVYLDSFEPHAEGEREISVAEILLRKSKFPSPALPIWARPDPNPPLESWKDPT
VPVVGHCPLPTKAPIPPPRRKTVLTVLESNVSSALAEATKTTFSSSGSADSYG
ATALPDQASDDGKSDVESYNSMPLLEGPGDPLSDGWSVSEASRDVVCSSM
YTWGALTPTCAAESEKILPINLSNLRHNMVYATTSRSASIRKQVTFDFRLOVD
DHYRDLKEMKASTVKALLSTEEACKLTPHSAKSKFGYAKDVNRLSSRAVNI
RSWNEDLLETETPIDTINAKNEVFCVPEKGRKPARLIPFDLGRVCEKEMALID
VSTLPQAVMGSSYFGYSPKQPELVNTWKKCPMGFSYDTPDFDSTVETSDIRV
EESIYOCCLDAPARQAIRSLERLYIGGLPTNSKQNCNGYRRCRASGVLTTCGNL
TCYLKATAACRAKLOQDTMLVNGDDLVVICESAGTOEDAAALRVTEAMTSAPPG
DPQPEYDLLELITSCSNVSVAHDSKQVYLYLTDPTIPLARAAETAPHTPVNSKL
GNLIMAPALWABMILMTHEFSILLAQEQLDQCIYGACYSIEPLDLPQILERLH
GLSAFLSHSPGEINRVASCLRLGVPLRWRHARSVRKLLSGGGRAPTCGRYL
FNWAVRTKLKLTPIPAASQLDLSGWVAGYSGGDIYHSLSRARPRWFLCLLLLSVG
GIYLLPNR"

342..914

mat_peptide

mat_peptide
915..1490
/product="C protein"
mat_peptide
1491..2528
/product="E protein"
mat_peptide
2529..3359
/product="NS1/E2 protein"
mat_peptide
3360..5186
/product="NS2 protein"
mat_peptide
5187..6380
/product="NS3 protein"
mat_peptide
6005..6104
/product="NS4 protein"
stem_loop
6381..9371
/note="possible stem-loop structure in NS4 sequence"
mat_peptide
9413..9448
/product="NS5 protein"
misc_feature
1899 a 2828 c 2676 g 2045 t
/note="T-stretch of 36 bp"

Query Match 96.9% Score 690.6; DB 14; Length 9448;
Best Local Similarity 98.0%; Pred. No. 2.1e-166;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 gccagccctgatggggcgacactccacatagatcactccccctgtgaggaactactg 60
DB 1 GCCAGCCCGGATTTGGGGCGGACACTCCACCATAGTACTCCCTGTGAGNACTACTG 60
QY 51 tcttcacacagaaagctctagccatggcgtagtagtgtagtgtagtgtagtgtagtg 120
DB 51 TCTTCAACGACAGAAACGCTTAGCCATGCGTTAGTATGAGTGTGCTGAGCTCCAGGAC 120
QY 121 ccccccctccgggagagcattagtgctcggaacggtagtacacacgaattgccaag 180
DB 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGAATTCGCCAG 180
QY 131 gacacgggttccttttttgatcaacccgcctcaatgctgtaggatttggcggtgcccc 240
DB 131 GACACCGGTGCTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGCGTGCCTCC 240
QY 241 gcgagac-gctagcagtagtgctgggtcgcaaggccttggtactgctgtagtgctg 300
DB 241 CCGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCCCTTGGTACTGCTGATGATAGG 300
QY 301 gtgcttgagtgccccgggaggtctctgtagaccgtgcacatgagcacaaatcctaaac 360
DB 301 GTGCTTGGAGTGCCTCCGGGAGGTCTCTGAGACCGTGCACCATGAGCACGAATCCTAAAC 360
QY 361 ctcaagaaacacacacgtaacacacacacacacacacacacacacacacacacacac 420
DB 361 CTCAAGAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 420
QY 421 gtgctcagatcgttggtgagatttaccctgttcgcgcagggggcccgaggttggtgtgc 480
DB 421 GTGCTCAGATCGTGTGGAGTAGTGTCTGTTCGCCCGCAGGGGCCCGAGGTGTGGTGTGC 480
QY 481 gcgcagactaggaagacttcgcagcgggtcgcaacctctggaagcgcaacacacacacac 540
DB 481 GCGCGACTAGGAACACTTCCGAGCGGTGCAACCTCTGTTGATGCGGACAACTATATCCCA 540
QY 541 aggtcgc 600
DB 541 AGGCTGCCCGACCGAGGGAGGGCCCTGGGCTCAGCGCGGTGATCCCTTGGCGCCCTCTATG 600
QY 601 gcaagagggcagtggtggcagagtggtctctgtcgccccggcgcccgaggtcccgactatt 660
DB 601 GCAATGAGGCTTGGGGTGGCGAGAGTGGCTTCTGTCAACCCCGCGCTCCCGGCTTCTAGT 660
QY 661 ggggccccctggaccccccgctaggtctcgtaatttgggttaaggtcatcgat 713
DB 661 GGGCCCCACGGACCCCGCGTAGTTCGCTAATTTGGTAAAGTTCATCGAT 713

RESULT 11

AB049094

LOCUS

DEFINITION

Hepatitis C virus gene for polyprotein, complete cds,

isolate:HCVT161.

ACCESSION

AB049094

VERSION

AB049094.1

KEYWORDS

SOURCE

Hepatitis C virus (isolate:HCVT161, specific_host:human) serum cDNA

to genomic RNA.

ORGANISM

Hepatitis C virus

REFERENCE

AUTHORS

Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K.,

Hatahara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and

Mishiro, S.

TITLE

Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients

with hepatocellular carcinoma: the 'progression score' revisited

JOURNAL

REFERENCE

AUTHORS

Mishiro, S.

JOURNAL

Submitted (19-SEP-2000) to the DDBJ/EMBL/GenBank databases. Shunji

Mishiro, Toshiba General Hospital, Department of Medical Sciences;

6-3-22 Higashi Oh-1, Shinagawa-ku, Tokyo 140-8522, Japan

(E-mail: shunji.mishiro@po.toshiba.co.jp, Tel: 81-3-3764-8981,

Fax: 81-3-3764-8992)

FEATURES

1. 9480

/organism="Hepatitis C virus"

/isolate="HCVT161"

/specific_host="human"

/db_xref="taxon:11103"

/tissue_type="serum"

/note="HCV-genotype: 1b"

1..342

/note="complete"

343..9378

/codon_start=1

/product="polyprotein"

/protein_id="BAB18807.1"

/db_xref="GI:11559455"

/translation="MSTNPKPQKTKRNTNRRPQDKFPGGQIVGVVLLPRRQPLR

GVATRKTSERSQRRPQIPKARRPEGRWAQPGYPMPLYNGLGWAQWLLSPRG

SRPSGPTDPRRRNGLKVIDTLTCGFADLMGYPLVGAFLGVARALAHGVRYVED

GVNATGNLPCGSFIFLLALLSCLTIPASAYEVNRVSGVYHVTNDCNSSTIVYDAAD

VIMHPGCVPCVWEGNSRRCVVALPTLAARNSIPTTTIRHVDLLVGAAGFCAMV

VGDLCGSFVLSQLETFSPRRYQTVQDCNCSLYPGHVSCHRMWDMNWSPTTALV

SOLLRIQAVQVDMVAGAWGLAGLAYISMVGNWAKVILVLLPAGVDGRGTHVSGGQ

AGYNARGLSVLSFGASQRIQVNTNGSMHINRTALNCNDTLQTFGLASLFYVHKFNA

SGCPEMASCRPIDKFDQGGPITHVEPNI SDQRCYWHYAPRPGIYVPAQVQCPYV

CFTPSVYVNTDSSGVPTNMGNETOVLLNNTNRPPLGNWGFPMWNSGFTKTCG

GPPCNIGVGNNTLFCPTDCFRKHPEATYKCGSGPMLTPLCLVDYPRYLMHYPCTVN

PTIFKRVYVSGVEHRLSAACNWTBGERCDLVDRDRAELSPLLSTTBQWLPCSFPT

LPASTEGLHLHONVVDVQVLYGVSAVSVVWKEIVVLLFLLDARARVCAQWMLL

LVAQEAALKNVLLNLSAVAGTSTLSLPLVFFCAAWIKGLVPGAAYAFVGMWPLL

LILLALPPRAYADREMAASCGGVFGLALLTLSPHYKVELARLIMWLQYFTRAEAL

LIQWVPLPNVRGGDDAILITLCVAVHPELITFKILLAFGPIVMOVLQGLIRVYFV

RAQGLIRACMLVRAAGGHYVQMLKLAALTGVYVYDHLTPLRDWAHAGLRDLAVAV

EPVVSNNETKVIWGAADTAACGDDILSLPVSAARGREILLGLPADSFQGRWRLAPI

TAYSOQTRGLGCIITSLTGRDNQVEGEVQVSTATQSFLLATCVNGVCWTVYHAGS

KTLAGPKGFVMTYVNDODLVGWAQPPGARSLTPTCTGSSDLVLTVRHADYIPVRRR

GDSRGLSPRPVYVYKSGSGLPLGHAHVIFRAAVCTRGVAKAVDFIPVESNET

TMRSPVFDNSGPPAVPOTFOVHLHAPTGSCKSTKVPAAVAAOYKVLVLPNSVAAT

LGFAYMSKAGIDPNVRTITTGAPITVSTYKFLADGCSGAYDIICDECH

STDSITLIGLITVLDQETAGARLVVLTATPPGVSVPVPHNIEVALNTEIPEFYG

KAIDLEANGRHLLFCHSKKCDLAKLSGLGVNAVAYYRGLDVSVPYISGVDVVV

ATDALMTGYTDFSDICNTCVTDFSLDPTTIIETTVPODAVSRQRGRGTGR

GRGGIYRRTVPERGSPFDSVLCYECYDAGCAWYELTAPETSVRLRAYLNTPPVPC

QDHLFESVSTGLTHIDAHFLSQKQAGDNFPYLVAYQATVCARAQAPPPSWDMQW

CLIRLKPGLHGPPTLLYRLGAVQNEITLTHPTIKFIMACMSADLEVIVSTWVLVGVGL

AB049094 9480 bp RNA VRL 05-DEC-2000

Hepatitis C virus gene for polyprotein, complete cds,

isolate:HCVT161.

AB049094

AB049094.1

GI:11559454

Hepatitis C virus (isolate:HCVT161, specific_host:human) serum cDNA

to genomic RNA.

Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

Hepacivirus.

1 (sites)

Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K.,

Hatahara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and

Mishiro, S.

Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients

with hepatocellular carcinoma: the 'progression score' revisited

Unpublished

2 (bases 1 to 9480)

Mishiro, S.

Direct Submission

Submitted (19-SEP-2000) to the DDBJ/EMBL/GenBank databases. Shunji

Mishiro, Toshiba General Hospital, Department of Medical Sciences;

6-3-22 Higashi Oh-1, Shinagawa-ku, Tokyo 140-8522, Japan

(E-mail: shunji.mishiro@po.toshiba.co.jp, Tel: 81-3-3764-8981,

Fax: 81-3-3764-8992)

Location/Qualifiers

1. 9480

/organism="Hepatitis C virus"

/isolate="HCVT161"

/specific_host="human"

/db_xref="taxon:11103"

/tissue_type="serum"

/note="HCV-genotype: 1b"

1..342

/note="complete"

343..9378

/codon_start=1

/product="polyprotein"

/protein_id="BAB18807.1"

/db_xref="GI:11559455"

/translation="MSTNPKPQKTKRNTNRRPQDKFPGGQIVGVVLLPRRQPLR

GVATRKTSERSQRRPQIPKARRPEGRWAQPGYPMPLYNGLGWAQWLLSPRG

SRPSGPTDPRRRNGLKVIDTLTCGFADLMGYPLVGAFLGVARALAHGVRYVED

GVNATGNLPCGSFIFLLALLSCLTIPASAYEVNRVSGVYHVTNDCNSSTIVYDAAD

VIMHPGCVPCVWEGNSRRCVVALPTLAARNSIPTTTIRHVDLLVGAAGFCAMV

VGDLCGSFVLSQLETFSPRRYQTVQDCNCSLYPGHVSCHRMWDMNWSPTTALV

SOLLRIQAVQVDMVAGAWGLAGLAYISMVGNWAKVILVLLPAGVDGRGTHVSGGQ

AGYNARGLSVLSFGASQRIQVNTNGSMHINRTALNCNDTLQTFGLASLFYVHKFNA

SGCPEMASCRPIDKFDQGGPITHVEPNI SDQRCYWHYAPRPGIYVPAQVQCPYV

CFTPSVYVNTDSSGVPTNMGNETOVLLNNTNRPPLGNWGFPMWNSGFTKTCG

GPPCNIGVGNNTLFCPTDCFRKHPEATYKCGSGPMLTPLCLVDYPRYLMHYPCTVN

PTIFKRVYVSGVEHRLSAACNWTBGERCDLVDRDRAELSPLLSTTBQWLPCSFPT

LPASTEGLHLHONVVDVQVLYGVSAVSVVWKEIVVLLFLLDARARVCAQWMLL

LVAQEAALKNVLLNLSAVAGTSTLSLPLVFFCAAWIKGLVPGAAYAFVGMWPLL

LILLALPPRAYADREMAASCGGVFGLALLTLSPHYKVELARLIMWLQYFTRAEAL

LIQWVPLPNVRGGDDAILITLCVAVHPELITFKILLAFGPIVMOVLQGLIRVYFV

RAQGLIRACMLVRAAGGHYVQMLKLAALTGVYVYDHLTPLRDWAHAGLRDLAVAV

EPVVSNNETKVIWGAADTAACGDDILSLPVSAARGREILLGLPADSFQGRWRLAPI

TAYSOQTRGLGCIITSLTGRDNQVEGEVQVSTATQSFLLATCVNGVCWTVYHAGS

KTLAGPKGFVMTYVNDODLVGWAQPPGARSLTPTCTGSSDLVLTVRHADYIPVRRR

GDSRGLSPRPVYVYKSGSGLPLGHAHVIFRAAVCTRGVAKAVDFIPVESNET

TMRSPVFDNSGPPAVPOTFOVHLHAPTGSCKSTKVPAAVAAOYKVLVLPNSVAAT

LGFAYMSKAGIDPNVRTITTGAPITVSTYKFLADGCSGAYDIICDECH

STDSITLIGLITVLDQETAGARLVVLTATPPGVSVPVPHNIEVALNTEIPEFYG

KAIDLEANGRHLLFCHSKKCDLAKLSGLGVNAVAYYRGLDVSVPYISGVDVVV

ATDALMTGYTDFSDICNTCVTDFSLDPTTIIETTVPODAVSRQRGRGTGR

GRGGIYRRTVPERGSPFDSVLCYECYDAGCAWYELTAPETSVRLRAYLNTPPVPC

QDHLFESVSTGLTHIDAHFLSQKQAGDNFPYLVAYQATVCARAQAPPPSWDMQW

CLIRLKPGLHGPPTLLYRLGAVQNEITLTHPTIKFIMACMSADLEVIVSTWVLVGVGL

AAAYCLTTGSSVIVGRIILISGRPAIIPDBREVLYOEDEMEECASHLPYIEQMQLA
EFOKOKALGLOIATKQAEAPVVEVSKRWLEAFWAKHMNFISGITOLAGSLTLP
NPATSLMAFTASITSPITQTTQHTLNFNLGDMVAAQALAPPASAFAGTAGAAGVS
IGLKVLDILAGVAGVAGALVAFKVMGSDMPSTEDLVNLIIPALISGALVGVVCA
AILRRHVGPGGAGVQMMNRLTAFASRGNHVSPTHVVPESDAARATQILSLTITOLL
KRLHOWINEDCSTPCSGSWLRDVMWICITVLDRKTLQSKLLPRLPGVPFPCGCGY
RGVWRGDMOTPCGAQIAGHVKNGSMKLVGPKTCSNTWHGTPIINATYTGCTPS
PAPNYSRALWRVAEEYVEVTRVGEFHYTGTNTDNVVKPCVPAPAEFFELDQVRLH
RYAPACKPDLLEDEVTOGLNQYPVGSOCPEDVTVVTSMLDPPSHITAEARRK
LARGSPSSASSASQLSAPKATCTTHHSDPADLLEIANLLMQMGNTIRVESE
NKVILDSFDPLEAEDEREVSAEILRKTFFPAIPINWAPYNPPLDSWRDPO
YVPPVHCCPLPAPAKAPIPPPRKKTIVVLSBSTVSSVSLAEIATKTFSSVSSAVDSG
SATAPDOASDNGDTGSDVGSYSSMPLEGGPDPLSDGSHSTVSEEGEDVCCSM
SYTWTGALITPCAAEESKLPINALNSLLRHNNVYVTSRSASORQRKVFEDRLQVL
DDHYRDLVKMEKASTVKALLSVEEACKLTPHSKSKFGYGDVDRVNRVSRVNH
IRSVWKLDEDETETPIDTITMAKNFCVQPEKGRKPARLIVYDLDFGVRCSEKALY
DVSTLPQAVGSSSYFOYSPQORVFLVNAWMSKTKTFFAYDFRCDSTVTENDIR
VBEITYQSDLVPEAROAKSLTERLYIGGPLTNSKQNGCVRRRCASGLVITSCGT
LTCYLKASACRAAKLOCTMLVCCGDLVVICESAGTOEDASLRFVTEAMPYRNP
GDPPREYDELITSCSNVSAHADSKRYVYLRDPTTPLARAWETARHTPNSW
LGNIIIMPTLWARMILMTHFFSILLAEQEDKALDQIYGAISIEPLDLPOIIRL
HGLSFAFLSHSPGEINRVSCLRKLGPPLRWRHRSVRKLLSRGGRATCGKY
LFNVAWTKLKLTPIPAAASQLDLSGWFVAGYGGDIYHLSLRARPWFPMWCLLSVG
VGIVLLPNR"

9379..9415
/note="upstream of poly-U stretch"
9416..9480
/note="poly(U) length is indistinct
poly(U) stretch"
BASE COUNT 1906 a 2685 g 2071 t
ORIGIN

Query Match 96.9%; Score 690.6; DB 14; Length 9480;
Best Local Similarity 98.0%; Pred. No. 2.1e-166;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Oy 1 gccagccccctgattggggcgacactccaccatagatcaactccccctgtgaggaactactg 60
Db 2 GCCAGCCCCGATTGGGGCGACACTCCACCATTAGATCATCTCCCTGTGAGGAACACTG 61
Oy 61 ttctcagcagaaagcgttagcctggttagtgatgagtcgtagcagctcccgagc 120
Db 62 TCTTTCAGCAGAAAGCGCTAGCCATGGCGGTAGTATGATGAGTCTGCTGCAGCCTCAGGTC 121
Oy 121 cccccctccgggagagaccatagtgctcgggaacggtagacacacgggaattccag 180
Db 122 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCAG 181
Oy 181 gacgacggggtctcttcttgatcaaccgctcaatgcctggagatttggggcggtcccc 240
Db 182 GACGACCGGGTCTCTTCTTGGATCAACCCGCTCAATGCTTGGAGATTTGGCGGTGCCCC 241
Oy 241 gcgagactgtagcggagtagtggttggtcggaaagggccttggtgactgcctgtagg 300
Db 242 GCGAGACTGCTAGCCAGTAGTGTGGTTCGGAAGAGCGCTTGTGTACTGCTGTATAGG 301
Oy 301 gtgcttgcgagtgccccgggaggtctcgtagaccgtgcaccatgacgacaaatccataac 360
Db 302 GTGCTTCCGAGTGCCTCCCGGAGGTCTCGTAGACCGTGCATCATGACGACAAATCCATAAC 361
Oy 361 ctcaaaajaaaaacaaacgtaacacacacggcgccacacgagcgctcaagttccccggcg 420
Db 362 CTCAAAGAAAAACAAACGCTAACCAACCGCGCCACAGSAGCTTAAGTTCCCGGGCG 421
Oy 421 gtggtcagatcgttggtgagtttacctgttgcgcgacgggccccaggttgggtgtgc 480
Db 422 GTGGTTCAGATCGTGTGGTGGAGTTTACTGTGTCCGCGCAGCGGCCCCAGGTTGGGTGTC 481
Oy 481 gcgagactaggaagacttccgagcgggtcgcaacctggtgaaagcgacacacctatcccc 540
Db 482 GCSCGACTAGGAAGACTTCCGAGCGGTTCGCAACCTCGTGGGAAGCGGACACCACTATCCCCA 541

Search completed: March 1, 2002, 15:24:35
Job time: 37284 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 08:33:12 ; Search time 232.06 Seconds
(without alignments)
2634.116 Million cell updates/sec

Title: US-09-763-836-1
Perfect score: 713
Sequence: 1 gccagcccccgtatggggc.....atttgggtaagtcacgat 713

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	713	100.0	713	21	Hepatitis type C v
2	690.6	96.9	9595	20	Infectious hepatitis
3	690.6	96.9	9595	22	Nucleotide sequenc
4	690.6	96.9	9595	22	Infectious hepatitis
5	690.6	96.9	9595	22	Infectious hepatitis
6	689	96.6	780	21	Hepatitis C virus
7	689	96.6	1880	13	Hepatitis C virus
8	689	96.6	2540	13	Hepatitis C virus
9	689	96.6	2540	14	NANBV hepatitis vir
10	689	96.6	2540	15	NANBV genomic fra
11	687.4	96.4	2033	15	Hepatitis C Virus

12	687.4	96.4	2033	16	AAO86788	Hepatitis C virus
13	687.4	96.4	4987	15	AAO65322	Vaccinia virus pro
14	687.4	96.4	4987	16	AAO86799	Vaccinia virus vec
15	685.8	96.2	923	17	AAO28348	Hepatitis C virus
16	684.8	96.0	9472	14	AAO33282	Korean hepatitis C
17	682.6	95.7	7911	13	AAO32436	HCV antigen clone
18	682.4	95.7	9416	13	AAO21829	Non-A, non-B viral
19	682.4	95.7	9416	13	AAO20268	Non-A, non-B viral
20	682.4	95.7	9416	20	AAO59394	Non-A, non-B hepat
21	680.8	95.5	3360	17	AAO36777	Hepatitis C genome
22	680.8	95.5	9413	16	AAO3960	Partial HCV non-st
23	679.2	95.3	9413	16	AAO81559	Hepatitis C virus
24	679	95.2	9436	15	AAO63499	Blood transmissibl
25	676	94.8	9413	16	AAO80498	DNA encoding HCV p
26	670.4	94.0	1863	12	AAO15363	Fragment of NANB h
27	668.2	93.7	1554	13	AAO32451	HCV core-envelope
28	667.2	93.6	1734	14	AAO40429	Hepatitis C virus
29	665.6	93.4	1734	14	AAO40425	Hepatitis C virus
30	665.6	93.4	9405	14	AAO40426	Full-length Hepati
31	664	93.1	1734	14	AAO40428	Hepatitis C virus
32	664	93.1	1734	14	AAO40430	Hepatitis C virus
33	663.6	93.1	686	15	AAO44921	Hepatitis C virus
34	663.6	93.1	3461	15	AAO64068	Non-A, non-B hepat
35	663.6	93.1	3461	16	AAO30386	5'UTR/CORE/ENV/NS1
36	661.8	92.8	9401	17	AAO1882	Hepatitis C virus
37	660.8	92.7	1734	14	AAO40427	Hepatitis C virus
38	660.8	92.7	1734	14	AAO40436	Hepatitis C virus
39	660.2	92.6	9416	13	AAO22871	NANBV Hutch c59 is
40	660.2	92.6	9416	19	AAO59378	Hepatitis C virus
41	660	92.6	807	13	AAO20940	PCR-generated frag
42	659.2	92.5	1734	14	AAO40434	Hepatitis C virus
43	659.2	92.5	1734	14	AAO40435	Hepatitis C virus
44	658.8	92.4	3401	15	AAO64069	Non-A, non-B hepat
45	658.8	92.4	3401	16	AAO30387	5'UTR/CORE/ENV/NS1

ALIGNMENTS

RESULT 1
AAA08097 standard; cDNA; 713 BP.
XX
AC AAA08097;
XX
DT 22-JUN-2000 (first entry)
XX
XX Hepatitis type C virus nucleotide sequence SEQ ID NO:1.
XX
XX Hepatitis C virus; HCV; 5' UTR; 5'-nontranslational region; diagnosis;
XX gene expression; infection; IRES; viral; ss.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX 5'UTR 1..341
XX CDS 342..713
XX FT /*tag= a
XX FT /*tag= b
XX FT /*note= "no stop codon given"
XX
XX WC2000012691 AL.

XX
XX 69 MAR-2000.
XX
XX 08-JUL-1999; 99WO-JP03682.
XX
XX 27-AUG-1998; 98JP-0241367.
XX
XX (FUSO) FUSO PHARM IND LTD.
XX
XX Yamada O, Yoshida H, Zhang J;

DR WPI: 2000-237867/20.
 DR P-PSDB; AAT82358.
 XX Nucleic acid sequence which potentiates gene expression and contains
 PT 5'-non-translated region of a viral genome is incorporated in vectors
 PT for improving the expression of a useful gene -
 PT
 XX
 PS Claim 18; Page 82-84; 94pp; Japanese.
 XX
 CC The present invention describes an oligonucleotide sequence which
 CC potentiates the expression of a useful gene when incorporated in a
 CC gene expression vector, where the oligonucleotide sequence consists
 CC of the 5'-non-translated region (5'UTR) of a viral genome or its
 CC fragment or modified form. The 5'UTR sequence is useful in screening
 CC potential initiation factors of interaction with the viral IRES or
 CC IRES-dependency, and for treatment of diseases caused by cap-dependent
 CC mRNA translation, and in the diagnosis of the severity of HCV infection.
 CC The present sequence represents a Hepatitis type C virus (HCV) nucleotide
 CC sequence comprising a 5'UTR, which is used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 713 BP; 136 A; 218 C; 226 G; 133 T; 0 other;

Query Match 100.0%; Score 713; DB 21; Length 713;
 Best Local Similarity 100.0%; Pred. No. 3e-188;
 Matches 713; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gccagccccctgattggggcgacactccaccatagatcaactccctgtgaggaactactg 60
 Db 1 gccagccccctgattggggcgacactccaccatagatcaactccctgtgaggaactactg 60
 QY 61 tcttcacacagaaagcgtctagccatgagcgttagtatgagtgctgcagcctccaggac 120
 Db 61 tcttcacacagaaagcgtctagccatgagcgttagtatgagtgctgcagcctccaggac 120
 QY 121 ccccccctccggagagccatagtgctgcggaaccggtgagtagacacggaattgcag 180
 Db 121 ccccccctccggagagccatagtgctgcggaaccggtgagtagacacggaattgcag 180
 QY 181 gacgacccgggtctcttcttgatcaaacccgctcaatgctgagattggcggtgcccc 240
 Db 181 gacgacccgggtctcttcttgatcaaacccgctcaatgctgagattggcggtgcccc 240
 QY 241 gcgagactctagcgaagtagtggttcggaagccttggtgactgcctgtagagg 300
 Db 241 gcgagactctagcgaagtagtggttcggaagccttggtgactgcctgtagagg 300
 QY 301 gtgcttgagtgccccggggaggtctctgtagaccgtgaccatgagcacaaatcctaaac 360
 Db 301 gtgcttgagtgccccggggaggtctctgtagaccgtgaccatgagcacaaatcctaaac 360
 QY 361 ctcaaaagaaaaaacacgtaaacacacccgccccacagcagcgtcaagtcccgggcg 420
 Db 361 ctcaaaagaaaaaacacgtaaacacacccgccccacagcagcgtcaagtcccgggcg 420
 QY 421 gtggtcagatcggtggaggtttacctgttcgacgagggccccaggttggtgtgc 480
 Db 421 gtggtcagatcggtggaggtttacctgttcgacgagggccccaggttggtgtgc 480
 QY 481 gcgcgactaggaaactccgagcgggtcgcaacctgtggaagcgcaaacctatcccca 540
 Db 481 gcgcgactaggaaactccgagcgggtcgcaacctgtggaagcgcaaacctatcccca 540
 QY 541 aggtcgcggccgagggcaggaacctgggtcagcccggtatccttggccccctctatg 600
 Db 541 aggtcgcggccgagggcaggaacctgggtcagcccggtatccttggccccctctatg 600
 QY 601 gcaacagggcattgggtggcagatggctcctgtgcgcccgaggtcccgccctagtt 660
 Db 601 gcaacagggcattgggtggcagatggctcctgtgcgcccgaggtcccgccctagtt 660
 QY 661 gggggcccttcggaccccccgcgtaggtcgctgtaattgggtaagggtcatcgat 713

Db 661 gggggcccttcggaccccccgcgtaggtcgctgtaattgggtaagggtcatcgat 713
 RESULT 2
 AAX24843
 ID AAX24843 standard; DNA; 9595 BP.
 XX
 AC AAX24843;
 XX
 DT 21-JUN-1999 (first entry)
 XX
 DE Infectious hepatitis C virus genotype 1b strain HC-J4 genome.
 XX
 KW HCV; infectious clone; infection; diagnosis; therapy; vaccine;
 XX screening; assay; antiviral; virucide; ss.
 KW
 CS Hepatitis C virus.
 XX
 FH Key: Location/Qualifiers
 FT CDS 342..9374
 FT /*tag= a
 XX
 PN WO9904008-A2.
 XX
 PD 28-JAN-1999.
 XX
 PF 16-JUL-1998; 98WO-US14688.
 XX
 PR 27-JAN-1998; 98US-0014416.
 PR 18-JUL-1997; 97US-0053062.
 XX
 FA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 FI Ruth J, Emerson SU, Purcell RH, Yanagi M;
 DR WP.; 1999-132252/11.
 DR P-PSDB; AAW98022.
 XX
 PT NC-1 isolated hepatitis C virus nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of HCV
 PT infections and for developing screening assays
 XX
 PS Claim 3; Fig 14A-F; 126pp; English.
 CC
 CC The present sequence comprises the nucleic acid sequence of the
 CC genome of infectious hepatitis C virus (HCV) genotype 1b strain
 CC HC-J4 (ATCC 209596) that is capable of expressing this virus when
 CC transfected into cells. HC-J4 was obtained from acute phase plasma
 CC of a chimpanzee experimentally infected with serum containing
 CC HC-J4/91. The claimed infectious nucleic acid sequence can be used
 CC to produce chimeric genomes (see AAX24833) consisting of the open
 CC reading frames of infectious nucleic acid sequences of other
 CC genotypes (including genotypes 1-6) and subtypes (such as 1b, 2a,
 CC 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also relates to
 CC the introduction of mutations or deletions into infectious nucleic
 CC acid sequences in order to produce an attenuated HCV virus suitable
 CC for vaccine development. Infectious nucleic acid sequences can
 CC also be used to produce attenuated virus via passage in vitro or in
 CC vivo of the viruses produced by transfection of a host cell with
 CC the infectious nucleic acid sequence. Vaccines comprising one or
 CC more polypeptides made from the infectious nucleic acid sequence are
 CC used to immunise mammals, especially humans, against hepatitis C.
 CC The nucleic acid sequences can also be used to induce protective
 CC immunity against the virus. The nucleic acid sequences or their
 CC encoded proteases (e.g. NS3 protease) can additionally be used to
 CC develop screening assays to identify antiviral agents for HCV.
 XX
 SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;

Query Match 96.9%; Score 690.6; DB 20; Length 9595;
 Best Local Similarity 98.0%; Pred. No. 1e-181;

Db 61 tcttcacgcagaaacgctctagccatggcgttagttagtatgagtgtctgtgagcctccagac 120
 QY 121 ccccccctccggagagccatagtggtctgcggaacccggtgagtacacgcgaattgccag 180
 Db 121 ccccccctccggagagccatagtggtctgcggaacccggtgagtacacgcgaattgccag 180
 QY 181 gacgacccggctctcttcttgatcaacccgctcaatgctgagatttggcgtgccccc 240
 Db 181 gacgacccggctctcttcttgatcaacccgctcaatgctgagatttggcgtgccccc 240
 QY 241 gcgagactgtagccgagtagtggttgcggaagcccttggtactgctgctgattagg 300
 Db 241 gcgagactgtagccgagtagtggttgcggaagcccttggtactgctgctgattagg 300
 QY 301 gtgcttgcagtgccccgggaggtctctgtagaccgtgcacatgagcaaatcctaac 360
 Db 301 gtgcttgcagtgccccgggaggtctctgtagaccgtgcacatgagcaaatcctaac 360
 QY 361 ctcaaaagaaaaaacgtaacacaaacccgcccagagcagcgtcaagttcccggcg 420
 Db 361 ctcaaaagaaaaaacgtaacacaaacccgcccagagcagcgtcaagttcccggcg 420
 QY 421 gtggtcagatcgtgtgtagttacctgttgcgcgcgagggcccaggttgggtgtgc 480
 Db 421 gtggtcagatcgtgtgtagttacctgttgcgcgcgagggcccaggttgggtgtgc 480
 QY 481 gcgcgactagaagacttccgagcgggtgcgaacctctgtggaagcgacaacctatcccca 540
 Db 481 gcgcgactagaagacttccgagcgggtgcgaacctctgtggaagcgacaacctatcccca 540
 QY 541 aggtctccggccgagagccgagcctgggtctagcccggtatcttggccctctatg 600
 Db 541 aggtctccggccgagagcctgggtctagcccggtatcttggccctctatg 600
 QY 601 gcaacgagggcatgggttggcagagatgctcctgtcccgccggtcccgccgctagt 660
 Db 601 gcaacgagggcatgggttggcagagatgctcctgtcaacctgtcaacctgggtcccgccgctagt 660
 .QY 661 ggggccccttcggaccccccgcgtaggtgcgtcaatttgggtaaggtcatgat 713
 Db 661 ggggccccttcggaccccccgcgtaggtgcgtcaatttgggtaaggtcatgat 713

RESULT 7

AAQ24467

ID AAQ24467 standard; DNA; 1880 BP.

XX AC AAQ24467;

XX DT 09-NOV-1992 (first entry)

XX DE NANB hepatitis virus strain HC-J4 genome.

XX DE non-A, non-B hepatitis virus; NANBH; PCR; amplification

KW polymerase chain reaction; vaccine; antibody; ss.

XX OS Non-A, non-B hepatitis virus.

XX FT Key Location/Qualifiers

FT CDS 342..1880

FT FT /*tag= a

FT FT /label= HC-J4

XX EP485209-A.

PN 13-MAY-1992.

XX PD 07-NOV-1991; 91EP-0310297.

XX PF 08-NOV-1990; 90JP-0304405.

XX PR (IMMO) IMMUNO JAPAN INC.

XX PA

XX XX

PI Nakamura T, Okamoto H;

XX WPI: 1992-160959/20.

XX P-PSDB: AAR24087.

XX XX

PI Recombinant cDNA of NANBH virus strain HC-J5 and corresp.

XX peptides - useful for diagnosis and in vaccines and immunological

XX pharmaceuticals

XX Disclosure: Page 11; 42pp; English.

XX This sequence is the genome of the non-A, non-B hepatitis virus
 CC (NANBH) strain HC-J4. This sequence was derived by amplification
 CC by polymerase chain reaction. The nucleotide sequences derived from
 CC this amplification can be used to detect NANBH infection which could
 CC not be detected by conventional methods. The detection kits allow
 CC highly specific and sensitive detection at an early phase of
 CC infection. The polypeptide product of this coding sequence can be used
 CC for the manufacture of vaccines and immunological pharmaceuticals
 CC and also to produce antibodies specific to NANBH.

XX SQ Sequence 1880 BP; 333 A; 593 C; 556 G; 398 T; 0 other;

Query Match 96.6%; Score 689; DB 13; Length 1880;

Best Local Similarity 97.9%; Pred. No. 1.8e-181;

Matches 698; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 gccagccccctgattggggcgacactccacatagatcacctccctgtgagaaactactg 60
 Db 1 gccagccccctgattggggcgacactccacatagatcacctccctgtgagaaactactg 60

QY 61 tcttcacgcagaaacgctctagccatggctgtagttagtgcgtgcagctccagac 120
 Db 61 tcttcacgcagaaacgctctagccatggctgtagttagtgcgtgcagctccagac 120

QY 121 ccccccctccggagagccatagtggtctgcggaacccggtgagtaacacgaattgccag 180
 Db 121 ccccccctccggagagccatagtggtctgcggaacccggtgagtaacacgaattgccag 180

QY 131 gacgacccgggtcccttcttgatcaacccgctcaatgctgagatttggcgtgccccc 240
 Db 131 gacgacccgggtcccttcttgatcaacccgctcaatgctgagatttggcgtgccccc 240

QY 241 gcgagactgtagccgagtagtggttgcggaagccctgtggtactgcctgattagg 300
 Db 241 gcgagactgtagccgagtagtggttgcggaagccctgtggtactgcctgattagg 300

QY 301 gtgcttgcagtgccccgggaggtctctgtagaccgtgcacatgagcaaatcctaacc 360
 Db 301 gtgcttgcagtgccccgggaggtctctgtagaccgtgcacatgagcaaatcctaacc 360

QY 361 ctcaaaagaaaaaacgtaacacaaacccgcccagagcagcgtcaagttcccggcg 420
 Db 361 ctcaaaagaaaaaacgtaacacaaacccgcccagagcagcgtcaagttcccggcg 420

QY 421 gtggtcagatcgttgggtggagttacctgttgcgcgcgagggcccaggttgggtgtgc 480
 Db 421 gtggtcagatcgttgggtggagttacctgttgcgcgcgagggcccaggttgggtgtgc 480

QY 481 gcgcgactagaagacttccgagcgggtcgcaacctctgtggaagcgcaacatcctacca 540
 Db 481 gcgcgactagaagacttccgagcgggtcgcaacctctgtggaagcgcaacatcctacca 540

QY 541 aggtctccggcccgagggcagggcagcctgggtcagccggtatccttggccctctatg 600
 Db 541 aggtctccggcccgagggcagggcagcctgggtcagccggtatccttggccctctatg 600

QY 601 gcaacgagggcatgggttggcagagatgctcctgtcccgccggtcccgccgctagt 660
 Db 601 gcaacgagggcatgggttggcagagatgctcctgtcaacctgtcaacctgggtcccgccgctagt 660

QY 661 ggggccccttcggaccccccgcgtaggtgcgtcaatttgggtaaggtcatgat 713


```

Db 661 ggggccccagaccgccggcgtaggtcgtaactgggtaagtcacgat 713
|||||
RESULT 8
AAQ29628
ID AAQ29628 standard; DNA; 2540 BP.
XX AC AAQ29628;
XX DT 16-MAR-1993 (first entry)
XX DE Hepatitis C virus HC-J4 5' region.
XX KW Non-A non-B hepatitis; NANBH; HCV; detection; diagnosis; screening;
XX KW PCR; primer; polymerase chain reaction; ss.
XX OS Hepatitis C virus.
XX PN EP510952-A.
XX PD 28-OCT-1992.
XX PF 23-APR-1992; 92EP-0303625.
XX PR 26-APR-1991; 91JP-0191376.
XX PA (IMMO ) IMMUNO JAPAN INC.
XX PI Nakamura T, Okamoto H;
XX DR WPI; 1992-359137/44.
XX PT Detection of non-A, non-B hepatitis virus - using new
XX PT oligo-nucleotide primers with nucleotide sequences corresp. to
XX PT part. of the viral RNA
XX PS Disclosure; Page 18; 54pp; English.
XX CC This sequence represents the 5' region of hepatitis C virus RNA. The
CC: original sample was obtained from human and chimpanzee plasma. RNA
CC was isolated from several samples and homology compared, and the
CC respective sequence of about 1900 - 2500 nucleotides of the 5'
CC terminus and 1100 nucleotides of the 3' terminus determined. The 5'
CC region (given) contains a non-coding region of at least 340
CC nucleotides and a region coding for the structural protein followed
CC by a region coding for the non-structural protein (none actually
CC detailed on the sequence given in the specification). When compared
CC with the sequence of HCV disclosed in EP-388232 this sequence showed
CC homology of 80.5%.
XX SQ Sequence 2540 BP; 472 A; 775 C; 741 G; 552 T; 0 other;

Query Match 96.6%; Score 689; DB 13; Length 2540;
Best Local Similarity 97.9%; Pred. No. 1.9e-181;
Matches 698; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 gccagcccccctgatggggcgacactccaccatagatcacctcccctgtgaggaaactactg 60
Db |||||||
Db 1 gccagcccccctgatggggcgacactccaccatagatcacctcccctgtgaggaaactactg 60
QY 61 tcttcacgcagaaaacgctctagccatgcggttagtatgtagtgcgagcctccaggac 120
Db |||||||
Db 61 tcttcacgcagaaaacgctctagccatgcggttagtatgtagtgcgagcctccaggac 120
QY 121 ccccccctccggagagaccatagtgctgcggaacgggtgagtacacggaaatgccag 180
Db |||||||
Db 121 ccccccctccggagagaccatagtgctgcggaacgggtgagtacacggaaatgccag 180
QY 181 gacgacgggtctcttcttggatcaaccgcgtcaatgctctgagattggcggtgccccc 240
Db |||||||
Db 181 gacgacgggtctcttcttggatcaaccgcgtcaatgctctgagattggcggtgccccc 240

```

```

241 gcgagactgctagccgagtagtgggtcgcgaaaggcccttggtagctgctgatagg 300
|||||
241 gcgagactgctagccgagtagtgggtcgcgaaaggcccttggtagctgctgatagg 300
C7 301 gtgcttgagtgccccgggaggtctctgtagaccctgacacatgacacaaatacctaaac 360
|||||
301 gtgcttgagtgccccgggaggtctctgtagaccctgacacatgacacaaatacctaaac 360
QY 361 ctcaagaaaaaacacgttaacacacccgcccacagagacgtcaagtccccggggcg 420
|||||
361 ctcaagaaaaaacacgttaacacacccgcccacagagacgtcaagtccccggggcg 420
QY 421 gtggtcagatcttggtggagtttacctgttgcgcgaggggccccaggttgggtgtgc 480
|||||
421 gtggtcagatcttggtggagtttacctgttgcgcgaggggccccaggttgggtgtgc 480
QY 481 gcgcgactagaagacttccgagcgtgcgaacctgtggaagcgacaaacctatcccca 540
|||||
481 gcgcgactagaagacttccgagcgtgcgaacctgtggaagcgacaaacctatcccca 540
QY 541 aggtctgcgcggccgagggcaggaacctgggtcagccgggtatccttggccccctctatg 600
|||||
541 aggtctgcgcggccgagggcaggaacctgggtcagccgggtatccttggccccctctatg 600
Db 541 aggtctgcgcggccgagggcaggaacctgggtcagccgggtatccttggccccctctatg 600
QY 601 gcaacgagggcagtggtgggagggatggctcctctgcgcgcgggctccccggcctagtt 660
|||||
601 gcaatgagggcttggtgggagggatggctcctctgcgcgcgggctccccggcctagtt 660
QY 661 ggggcccccttcgacccccggcgtaggtcgcgtaatttgggtgaagtcacgat 713
|||||
661 ggggccccacgagccccggcgtaggtcgcgtaatttgggtgaagtcacgat 713

```

```

RESULT 9
AAQ43889
ID AAQ43889 standard; cDNA to RNA; 2540 BP.
XX AC AAQ43889;
XX DT 21-OCT-1993 (first entry)
XX DE NANB hepatitis virus polynucleotide N-2540-2.
XX KW Non-A, non-B; virus; polymerase chain reaction; detection;
XX KW sensitive; specific; HCV; NANBH; ss.
XX OS Non-A, non-B hepatitis virus.
XX PH Key Location/Qualifiers
XX FT CDS 342..2540
XX FT 5'UTR /*tag= a
XX FT /*tag= b
XX FT /*note= "from 5' terminal of NANBH virus RNA"
XX JF JF05091884-A.
XX PD 16-APR-1993.
XX PF 10-APR-1991; 91JP-0196175.
XX ER 10-JUN-1990; 90JP-0153401.
XX ER 08-NOV-1990; 90JP-0304405.
XX PA (NAKA/) NAKAMURA T.
XX KW WPI; 1993-199637/25.
XX DR P-PSDB; AAR38279.
XX AN Antigen related to non-A and non-B hepatitis virus - comprises
XX PT non-translation region comprising 340 - 341 mols. of nucleotides,
XX PT non-translation region comprising 1885 - 2551 mols. of

```

PT nucleotides including region 1,149 and, etc.

PS Claim 3; Page 19-20; 73pp; Japanese.

XX The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
XX which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
CC The polypeptide it encodes may be used in a system for detecting
CC NANB hepatitis. This method is highly specific and sensitive, and
CC can detect NANB hepatitis virus which could not be detected by
CC conventional methods.

XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;

Query Match 96.6%; Score 689; DB 14; Length 2540;
Best Local Similarity 97.9%; Pred. No. 1.9e-181;
Matches 698; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
QY 1 gccagccccctgattggggcgacactccaccatagatcaactccctctgaggaaactactg 60
Dd 1 gccagccccctgattggggcgacactccaccatagatcaactccctctgaggaaactactg 60
QY 61 tcttcacgcagaaagcgtctagccatggcgttagatagtgctgcagcctccaggac 120
Dd 61 tcttcacgcagaaagcgtctagccatggcgttagatagtgctgcagcctccaggac 120
QY 121 ccccccctccgggagagccatagtggtctgcggaacccggtgagtacaccggaattgccag 180
Dd 121 ccccccctccgggagagccatagtggtctgcggaacccggtgagtacaccggaattgccag 180
QY 181 gacacacgggtccttttcttgatcaaacccgctcaatcctcgagatgtggcgtgccccc 240
Dd 181 gacacacgggtccttttcttgatcaaacccgctcaatcctcgagatgtggcgtgccccc 240
QY 241 gcgagactgctagcgcgagtagtggtggtcggaagcccttggttactgcctgatagg 300
Dd 241 gcgagactgctagcgcgagtagtggtggtcggaagcccttggttactgcctgatagg 300
QY 301 gtgttcgagtgccccggagggtctctgtagaccgtgcacatgagcacaaatcctaacc 360
Dd 301 gtgttcgagtgccccggagggtctctgtagaccgtgcacatgagcacaaatcctaacc 360
QY 361 ctcaagaaacaaacaaacgttaacacacccgcccacacgagcgtcaagtctccggggc 420
Dd 361 ctcaagaaacaaacaaacgttaacacacccgcccacacgagcgtcaagtctccggggc 420
QY 421 gtgttcagatcgttggtggagttacctgttgcgcgcagggccccaggttggtgtgtgc 480
Dd 421 gtgttcagatcgttggtggagttacctgttgcgcgcagggccccaggttggtgtgtgc 480
QY 481 gcgcgactaggaagacttccgagcgggtcgcaacctctggaagggcgacaacctatcccca 540
Dd 481 gcgcgactaggaagacttccgagcgggtcgcaacctctggaagggcgacaacctatcccca 540
QY 541 aggtctccggccgagggcaggaacctggctcagccgggtatccttgccccctctatg 600
Dd 541 aggtctccggccgagggcagggccctggctcagccgggtatccttgccccctctatg 600
QY 601 gcaacgagggcagtggtggcgaggaaggctcctgtgcgccccggtcccgccgctcagtt 660
Dd 601 gcaatgagggcttggtggcgaggaaggctcctgtgcaccaaccggtcccgccgctcagtt 660
QY 661 ggggccccctcgagccccccgctaggttcgcgtaatttggttaagggtatcgat 713
Dd 661 ggggccccacagggaccccccgctaggttcgcgtaacttggttaagggtatcgat 713
```

RESULT 10

AAQ63753

ID AAQ63753 standard; cDNA to genomic RNA; 2540 BP.

XX

AC AAQ63753;

XX

30-JAN-1995 (first entry)

XX NANBHV genomic fragment #2.

XX Polymerase chain reaction; PCR; primer; amplify; detection; NANBHV;
XX non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.
XX Synthetic.

XX JP6125777-A.

XX

PD 10-MAY-1994.

XX

20-JUN-1991; 91JP-0247120.

XX

20-JUN-1991; 91JP-0247120.

XX

XX (NAKA/) NAKAMURA T.

XX

XX WPI; 1994-187937/23.

XX

XX Oligonucleotide primer pairs specific for non-A, non-B hepatitis
XX virus - used for high sensitivity detection of non-A non-B (NANB)
XX hepatitis virus

XX Disclosure; Page 24-25; 25pp; Japanese.

XX The sequences given in AAQ63752-53 represent fragments of the non-A,
XX non-B hepatitis virus (NANBHV) genome. These fragments were amplified
XX using the primers given in AAQ63732-51. These primers were used in the
XX detection of NANBH. The primers are based on the 5'-terminal region and
XX the core protein coding region. The method allows highly sensitive
XX detection of NANBH.

XX Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 other;

Query Match 96.6%; Score 689; DB 15;

Best Local Similarity 97.9%; Pred. No. 1.9e-181;

Matches 698; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
QY 1 gccagccccctgattggggcgacactccaccatagatcaactccctctgaggaaactactg 60
Dd 1 gccagccccctgattggggcgacactccaccatagatcaactccctctgaggaaactactg 60
QY 61 tcttcacgcagaaagcgtctagccatggcgttagatagtgctgcagcctccaggac 120
Dd 61 tcttcacgcagaaagcgtctagccatggcgttagatagtgctgcagcctccaggac 120
QY 121 ccccccctccgggagagccatagtggtctgcggaacccggtgagtacaccggaattgccag 180
Dd 121 ccccccctccgggagagccatagtggtctgcggaacccggtgagtacaccggaattgccag 180
QY 181 gacacacgggtccttttcttgatcaaacccgctcaatcctcgagatgtggcgtgccccc 240
Dd 181 gacacacgggtccttttcttgatcaaacccgctcaatcctcgagatgtggcgtgccccc 240
QY 241 gcgagactgctagcgcgagtagtggtggtcggaagcccttggttactgcctgatagg 300
Dd 241 gcgagactgctagcgcgagtagtggtggtcggaagcccttggttactgcctgatagg 300
QY 301 gtgttcgagtgccccggagggtctctgtagaccgtgcacatgagcacaaatcctaacc 360
Dd 301 gtgttcgagtgccccggagggtctctgtagaccgtgcacatgagcacaaatcctaacc 360
QY 361 ctcaagaaacaaacaaacgttaacacacccgcccacacgagcgtcaagtctccggggc 420
Dd 361 ctcaagaaacaaacaaacgttaacacacccgcccacacgagcgtcaagtctccggggc 420
QY 421 gtgttcagatcgttggtggagttacctgttgcgcgcagggccccaggttggtgtgtgc 480
Dd 421 gtgttcagatcgttggtggagttacctgttgcgcgcagggccccaggttggtgtgtgc 480
```



```

XX 02 SEP-1993;.. 93JP-0241973.
XX (MITU ) MITSUBISHI KASEI CORP.
XX FA
XX 1995-144713/19.
XX DR
XX P: SDB; AAR72801.
XX
XX Antiviral agent comprising component which disrupts viral gene
XX translation; used for the selective inhibition of e.g. Hepatitis
XX C virus, polio:virus and human rhinovirus
XX
XX Disclosure: Pages 19-22; 23pp; Japanese.
XX
XX AA086788 encodes AAR72800 Hepatitis C virus (HCV) T7N1-19, which
XX disrupts viral gene translation, by preventing the binding of
XX the viral mRNA to the internal ribozyme entry site. It was
XX used in the construction of an antiviral agent detecting vector
XX (AA086799 which encodes AAR72801), which comprises a vaccinia virus
XX vector, a vector containing T7N1-19 and a firefly luciferase gene.
XX The antiviral agent can be used to treat HCV, poliovirus, cerebral
XX cardio-hepatitis, human rhinovirus and foot and mouth disease viral
XX infections.
XX
XX Sequence 4987 BP; 1516 A; 1050 C; 1081 G; 1322 T; 18 other:
XX
XX Query: Match 96.4%; Score 687.4; DB 16; Length 4987;
XX Best Local Similarity 97.8%; Pred. No. 6.6e-181;
XX Matches 697; Conservative 0; Mismatches 16; Indels 0; Gaps
XX
XX 1 gcagcccctgattggggcgacactccaccatagatcactcccctgtgaggaactactg 60
XX 1495 gccagcccctgattggggcgacactccaccatagatcactcccctgtgaggaactactg 1544
XX
XX 51 tcttcacgagaaagcgtctagccaatgacgttagtatgagtgcgtgcagcctccaggac 120
XX 1545 tcttcacgagaaagcgtctagccaatgacgttagtatgagtgcgtgcagcctccaggac 1604
XX 121 cccctctccggagagccatagtgtgtctgcggaaccggtgagtcacaccggaattgccag 180
XX 1605 cccctctccggagagccatagtgtgtctgcggaaccggtgagtcacaccggaattgccag 1664
XX 181 gacgaccgggtcctttcttgataccaaccgcgtccaatgctggagattggcggtgcctccc 240
XX 1665 gacgaccgggtcctttcttgataccaaccgcgtccaatgctggagattggcggtgcctccc 1724
XX 241 gcgagactgctagccgagtagtgttggtgcggaaggccttgggtactgctgatagg 300
XX 1725 gcgagac-gctagccgagtagtgttggtgcggaaggccttgggtactgctgatagg 1784
XX 301 gtgcttcgagtgccccggagggtctctgtagaccgtgcacatgagcacaaatcctaaac 360
XX 1785 gtgcttcgagtgccccggagggtctctgtagaccgtgcacatgagcacaaatcctaaac 1844
XX 361 ctcaagnaaaacacgaactaacacaaaccgcgccacagagcgtcaagtctcccgggcg 420
XX 1845 cccaaagaaatcaaacgtaaacacaaaccgcgccacagagcgttaagtctcccgggcg 1904
XX 421 gtgggtcagatcgttggtggagtttacctgtttgcgcgcagggggccccagggttggtgtgc 480
XX 1905 gtgggtcagatcgttggtggagtttacctgtttgcgcgcagggggccccagggttggtgtgc 1964
XX 481 gcgcgactaggaagactcccgagcggtgcgaacctctgtgaagcgagacacactatcccca 540
XX 1965 gcgcgactaggaagactcccgagcgcgcaacctctgtgaagcgagacacactatcccca 2024
XX 541 aggcctgcgcgcgcgagggcaggacctgggtctcagcccggtatccttgccctctatg 600
XX 2025 aggcctgcgcgcgcgaggggtgagggccttggtctcagcccggtgtaaccttggtccctctatg 2084
XX 601 gcaacgagggcatgggggtgggcaggatggctctgtgccccgagcgtcccgcgctaggt 660
XX

```

241	QY	gcgagactgctagccgagtagtggttggttcgcgaagagccttggtagctgcctgatagg	300
1725	DB	gcgagacigctagccgagtagtggttggttcgcgaagagccttggtagctgcctgatagg	1784
301	QY	gtgcttgcgagtgcgcccgagggtctctagacggtgcacccatgagcacaaatcctaacc	360
1785	DB	gtgcttgcgagtgcgcccgagggtctctagacggtgcacccatgagcacaaatcctaacc	1844
361	QY	ctcaagaanaaaccaaacgtaacacaaaccgcgccacaggaagctcaagttcccgggcg	420
1845	DB	cccaagaanaaaccaaacgtaacacaaaccgcgccacaggaagctcaagttcccgggcg	1904
421	QY	gtggctcaatcgttgtaggtttaccttggtgcgcgcagagggcccccaggttgaggtgc	480
1905	DB	gtggctcagatcgttgtaggtttaccttggtgcgcgcagagggcccccaggttgaggtgc	1964
541	QY	gcgcgactaggagacttccgagcgttcgcaacctcgttggaagcgacaaacctatcccca	540
1965	DB	gcgcgactaggagacttccgagcgttcgcaacctcgttggaagcgacaaacctatcccca	2024
541	QY	aggcttcgcgcgccgagggcaggacctgggtctcagcccggtatccttggccctctatg	600
2025	DB	aggcttcgcgcgccgagggcaggacctgggtctcagcccggtatccttggccctctatg	2084
601	QY	gcaacgagggcattgggttgggcaggatggctctgtgcccccgaggtcccgcgctagtt	660

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 04:58:52 ; Search time 8196.34 Seconds
(without alignments)
934.776 Million cell updates/sec

Title: US-09-763-836-1

Perfect score: 713

Sequence: 1 gccagccctgtatggggc.....atttggtgaagtcacgat 713

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:*
- 1: em_estfun:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estom:*
 - 5: em_estpl:*
 - 6: em_estba:*
 - 7: em_estro:*
 - 8: em_estov:*
 - 9: em_htc:*
 - 10: gb_est1:*
 - 11: gb_est2:*
 - 12: gb_htc:*
 - 13: gb_gss:*
 - 14: em_gss_fun:*
 - 15: em_gss_hum:*
 - 16: em_gss_inv:*
 - 17: em_gss_pln:*
 - 18: em_gss_pro:*
 - 19: em_gss_rod:*
 - 20: em_gss_vrt:*
 - 21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	7.7	1101	13	CNS016UE
2	53.8	7.5	1201	13	CNS016BR
3	51.8	7.3	879	13	CNS02QOG
4	48.4	6.8	925	13	CNS00BNG
5	48.2	6.7	925	13	CNS0091P
6	47.6	6.7	925	13	CNS0091P
7	46.4	6.5	1036	13	CNS010BS
8	44.2	6.2	518	13	CNS016UP
9	44.2	6.2	932	13	CNS0072Q
10	43	6.0	844	13	CNS0052P
11	42.6	6.0	442	10	AF367693
12	42	5.9	932	13	CNS0072Q

C	13	40.4	5.7	384	13	CNS043G6	AL272751 Tetraodon
C	14	40.4	5.7	1203	13	CNS015Y4	AL106054 Drosophil
C	15	40.2	5.6	935	13	CNS006XK	AL106051 Drosophil
C	16	40	5.6	936	11	BG852371	BG852371 1024034A0
C	17	39.6	5.6	487	13	A2049537	A2049537 GSSBnu010
C	18	39.6	5.6	844	13	CNS0052P	AL056652 Drosophil
C	19	39.6	5.6	1009	13	CNS010EW	AL098882 Drosophil
C	20	39.6	5.6	1036	13	CNS010BS	AL098770 Drosophil
C	21	39.6	5.6	1201	13	CNS016BR	AL106545 Drosophil
C	22	39.2	5.5	961	13	CNS01087	AL098641 Drosophil
C	23	39	5.5	1201	13	CNS014BJ	AL103945 Drosophil
C	24	38.8	5.4	977	13	CNS00JX7	AL076850 Drosophil
C	25	38.4	5.4	870	13	CNS006EZ	AL064271 Drosophil
C	26	38	5.3	910	13	CNS006ON	AL065629 Drosophil
C	27	37.8	5.3	1101	13	CNS017RT	AL108419 Drosophil
C	28	37.4	5.2	1101	10	AL513871	AL513871 AL513871
C	29	37.2	5.2	491	13	CNS04QTL	AL303042 Tetraodon
C	30	37.2	5.2	685	11	BG282216	BG282216 602403372
C	31	37.2	5.2	1101	13	CNS017SY	AL108460 Drosophil
C	32	37	5.2	692	13	CNS007WH	AL050923 Drosophil
C	33	37	5.2	723	11	BE967041	BE967041 601660773
C	34	37	5.2	766	13	A2197732	A2197732 SP_1035_B
C	35	36.8	5.2	775	13	A2184156	A2184156 SP_1002_B
C	36	36.8	5.2	879	13	CNS02QOG	AL209545 Tetraodon
C	37	36.8	5.2	997	13	CNS006DN	AL065132 Drosophil
C	38	36.8	5.2	1009	13	CNS010EW	AL098882 Drosophil
C	39	36.8	5.2	1101	13	CNS017SY	AL108460 Drosophil
C	40	36.8	5.2	1143	13	AQ897991	AQ897991 HS_3122_B
C	41	36.6	5.1	366	11	R87116	R87116 YG30H12.s1
C	42	36.6	5.1	1098	13	CNS03F4J	AL241228 Tetraodon
C	43	36.4	5.1	1084	13	CNS05NAK	AL345125 Tetraodon
C	44	36.4	5.1	1100	13	CNS016KD	AL106855 Drosophil
C	45	36.4	5.1	1201	13	CNS01668	AL106346 Drosophil

ALIGNMENTS

RESULT 1

CNS015UE

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CNS015UE 1101 bp DNA GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence T7 end of BAC
BACNL6G16 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL107216
AL107216.1 GI:5626210
GSS.
fruit fly.
Plasmid Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91008 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACNL6G16"


```
BASE COUNT 67 a 317 c 213 g 99 t 183 others
ORIGIN
Query Match 7.3%; Score 51.8; DB 13; Length 879;
Best Local Similarity 41.0%; Pred. No. 0.0052;
Matches 128; Conservative 28; Mismatches 156; Indels 0; Gaps 0;
QY 379 gtaacacaccccccacagagcgtcaagttcccgggcggtgtgcagatcgttggtg 438
Db 503 GYBCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 562
QY 439 gaggttacctgttcgcgcagggggcccaaggttggtgtgcgcgactagagaactt 498
Db 563 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 622
QY 499 ccgagcgtgcgaacctcgttggaagcgacaaacctatcccccaaggctcgcgcgcgagg 558
Db 623 CCCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 682
QY 559 gcaggacctgggttcagccgggtatccttgccctctatgcaacaggggcatgggtt 618
Db 683 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 742
QY 619 ggcagagtggtcctgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 678
Db 743 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 802
QY 679 ggcgtaggtgcg 690
Db 803 GCGGGGGGGGGC 814

RESULT 4
CNS000BNG 1101 bp DNA GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC24B13 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL057398
VERSION AL057398.1 GI:4937864
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
Genoscope.
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Source 1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="BACR19D16"

BASE COUNT 129 a 276 c 132 g 151 t 413 others
ORIGIN
Query Match 6.8%; Score 48.4; DB 13; Length 1101;
Best Local Similarity 25.0%; Pred. No. 0.04;
Matches 62; Conservative 79; Mismatches 107; Indels 0; Gaps 0;
QY 460 gggggcccggttggtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 519
Db 701 GGMGMGGMGGRGKMGKMGNGMGGCGKSGKMGCGGCGKCKKKNKTKCKCKCKKK 760
QY 520 gaaggcgacaacctatcccccaaggctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 579
Db 731 CCGKNCKHCBCCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCK 820
QY 580 ggtatcctggccctctatggaacagagggcatgggggtggggcagatgctcctcgc 639
Db 821 CYTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 880
QY 640 cccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 699
Db 881 YCFBKBCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 940
QY 700 gtaaggtc 707
Db 941 CKYKGKNC 948

RESULT 5
CNS0091P 925 bp DNA GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL053013
VERSION AL053013.1 GI:4934461
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
Genoscope.
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Source 1. .925
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="BACR19D16"
```

BASE COUNT	120 a	61 c	61 g	172 t	511 others
ORIGIN	/note="end : TET3"				
Query Match	6.8% ; Score 48.2; DB 13; Length 925;				
Best Local Similarity	13.9% ; Pred No. 0.044; Mismatches 121; Indels 0; Gaps 0;				
MATCHES					
Qy	469	ggttggtgctgcgcgactaggaaagacttcgcgagcgatcgcgaacctctgtggaaggcac	528		
Dd	557	GSGYKKGCSGGBSCCSCSCSSSSCCBCBCCCCCSCSYCCSSBSKCSSTBS	616		
Qy	529	aacatcccccaagctccgcggccgagcaggcagaacctggctcagccgggttatcott	588		
Dd	617	CSCCSSKSVCGTSCSSSSSCSSSSTFSSSTSSSKSSSSSSSSSSSYTTKS	676		
Qy	589	ggcccctctatggcaacgagggctgggtggcgaggtgcctctcgcgcccgcgct	648		
Dd	677	ASGGSWAGCGSGTGTSSTSSSSSSTNSTSSVSCKSTBSSGCSBSSGSSSSST	736		
Qy	649	ccggcctagtgtggggcccttcgcaccgcccgtagtgcgtaatttgggtgaagtc	708		
Dd	737	SBBSCSTSTSSSSSSVSSSTCSCTCCCYSYSSSTSSWGTSGGSSSVGTSS	796		
Qy	709	tcat 713			
Dd	797	SSDST 801			
RESULT	6				
LOCUS	CNS0091P 925 bp DNA GSS 03-JUN-1999				
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL053013				
VERSION	AL053013.1 GI:4934461				
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 925)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (02-JUN-1999); Genoscope - Centre National de Sequençage :				
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr) - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm. Location/Qualifiers 1 . 925 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="RPCI-98" /clone="BACR19D16" /note="end : TET3"				
FEATURES					
source					
BASE COUNT	120 a	61 c	61 g	172 t	511 others

[illegible]

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)


```
Db 301 ggcgttcgagagcccgagaggtctcgtacagctgcacacagacgaatccctaac 360
Oy 361 ctcaagaataaaccaacgtaacacacacacgcccgcacacagagacgtatcccgagc 420
Db 361 ctcaagaataaaccaacgtaacacacacacgcccgcacacagagacgtatcccgagc 420
Oy 421 ggggtcagatctgtgtgagtttacctctgtccgcgcaggggcccagaattgggtgtc 480
Db 421 ggggtcagatctgtgtgagtttacctctgtccgcgcaggggcccagaattgggtgtc 480
Oy 481 ggcgcactaggaagacttcgcagcggtcgaacacctgtgaaggcgacacactatccca 540
Db 481 ggcgcactaggaagacttcgcagcggtcgaacacctgtgaaggcgacacactatccca 540
Oy 541 aggttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Db 541 aggttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Oy 601 gcaacgagagcattgggtgagagatgtctctgtccgcgcgcgcgcgcgcgcgcgcgc 660
Db 601 gcaacgagagcattgggtgagagatgtctctgtccgcgcgcgcgcgcgcgcgcgcgc 660
Oy 661 ggggccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
Db 661 ggggccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
```

```
RESULT 2
US-09-014-416-6
: Sequence 6, Application US/09014416
: Patent No. 6153421
: GENERAL INFORMATION:
: APPLICANT: Yanagii, Masayuki
: APPLICANT: Buhn, Jens
: APPLICANT: Emerson, Susanne U.
: APPLICANT: Purcell, Robert H.
: TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
: FILE REFERENCE: 20264276
: CURRENT APPLICATION NUMBER: US/09/014.416
: CURRENT FILING DATE: 1998-01-27
: EARLIER APPLICATION NUMBER: US 60/053.062
: EARLIER FILING DATE: 1997-07-18
: NUMBER OF SEQ ID NOS: 65
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 6
: LENGTH: 9599
: TYPE: DNA
: ORGANISM: Hepatitis C virus
US-09-014-416-6
```

```
Query Match 96.98; Score 690.6; DB 3; Length 9599;
Best Local Similarity 98.0%; Pred. NO. 7e-195;
Matches 699; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
Oy 1 gccacccccgatgggggggacactccacatagatcactccctgtggaactactg 60
Db 1 gccacccccgatgggggggacactccacatagatcactccctgtggaactactg 60
Oy 61 tcttcacgagaaagcgtacgcatggtatgtagtctgtgagcctccagagc 120
Db 61 tcttcacgagaaagcgtacgcatggtatgtagtctgtgagcctccagagc 120
Oy 121 cccccctccggagagcagatagtgctgcggaacgggtgagtaacgggaattgcag 180
Db 121 cccccctccggagagcagatagtgctgcggaacgggtgagtaacgggaattgcag 180
Oy 181 gacgacgggtctctcttgatcaaccgcgtcaatgcttggaattgggcggtcccc 240
Db 181 gacgacgggtctctcttgatcaaccgcgtcaatgcttggaattgggcggtcccc 240
```

```
Oy 241 ggcgacgtctagccgagtagtctgtgtgcgcgaagacgtctgtgtactgctgaag 300
Db 241 ggcgacgtctagccgagtagtctgtgtgcgcgaagacgtctgtgtactgctgaag 300
Oy 301 ggcgttcgagagcccgagaggtctcgtagacacgtgcacatgaacacaaatcccaac 360
Db 301 ggcgttcgagagcccgagaggtctcgtagacacgtgcacatgaacacaaatcccaac 360
Oy 361 ctcaagaataaaccaacgtaacacacacacgcccgcacacagagacgtatcccgagc 420
Db 361 ctcaagaataaaccaacgtaacacacacacgcccgcacacagagacgtatcccgagc 420
Oy 421 ggggtcagatctgtgtgagtttacctctgtccgcgcaggggcccagaattgggtgtc 480
Db 421 ggggtcagatctgtgtgagtttacctctgtccgcgcaggggcccagaattgggtgtc 480
Oy 481 ggcgcactaggaagacttcgcagcggtcgaacacctgtgaaggcgacacactatccca 540
Db 481 ggcgcactaggaagacttcgcagcggtcgaacacctgtgaaggcgacacactatccca 540
Oy 541 aggttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Db 541 aggttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Oy 601 gcaacgagagcattgggtgagagatgtctctgtccgcgcgcgcgcgcgcgcgcgcgc 660
Db 601 gcaacgagagcattgggtgagagatgtctctgtccgcgcgcgcgcgcgcgcgcgcgc 660
Oy 661 ggggccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
Db 661 ggggccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 713
```

```
RESULT 3
US-08-474-700B-45
: Sequence 45, Application US/08474700B
: Patent No. 6001990
: GENERAL INFORMATION:
: APPLICANT: Wands, Jack
: APPLICANT: Wakita, Takaji
: APPLICANT: Moradpour, Darius
: TITLE OF INVENTION: ANTISENSE INHIBITION OF HEPATITIS C
: TITLE OF INVENTION: VIRUS
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: COMPUTER: IBM PS/2 Model 502 or 55SX
: OPERATING SYSTEM: MS-DOS (Version 5.0)
: SOFTWARE: WordPerfect (Version 5.1)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474.700B
: FILING DATE: 07-JUN-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/240.382
: FILING DATE: 10 May 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Fraser, Janis K.
: REGISTRATION NUMBER: 34,819
: REFERENCE/DOCKET NUMBER: 00786/279001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 780 base pairs
```


STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333..1499
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..1499
OTHER INFORMATION: /note="sequence = 1 - 1499 of SEQ
OTHER INFORMATION: ID NO: 1"
US-08-324-977-3

Query Match 95.7%; Score 682.4; DB 1; Length 1499;
Best Local Similarity 98.4%; Pred. No. 8.8e-193;
Matches 689; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 14 tgggggagacactccaccatagatacctccctgtgaggaactacgtgtcaccgagaa 73
DB 5 tgggggggacgtccacacatagatcactccctgtgaggaactacgtgtcaccgagaa 64
QY 74 agcgtcagcagcgtgagtagtgcgtgcagccctccaggaaccccccctccggg 133
DB 65 AGGCTAGACCATGGCGGTAGTATGATGTCGTGACACCCCAAGAGACCCCCCTCCGGG 124
QY 134 agagccatagtggtctgcggaacgggtgagtaaccggaaattccagagacacggg 193
DB 125 AGAGCCATAGTGTCTGCGGAGACCGGTGAGTACACCCGGAATTGCCAAGAGACCGG 184
QY 194 ttcttgatcaaacccgctlaattgctcgagatttggcgctgcccccgagagctgtag 253
|||||

DB 185 TTCTTGATCAACCCGCTCATGCTGGAGATTGTGGCGCTGCCCGCAGACTGCTAG 214
QY 254 ccgagtagtctgtgtgcggaagccttgttgtaactcctgtagtgccttgcagtg 313
DB 245 CCGAGTAGTGTGGGTGGCGAAGGCCCTGTGTAGTGCCTGAPAGGCTGCTGGAGTG 304
QY 314 ccccggaaggtctcgtatagaccgtgacacatgaacaaatcctaactaaagaaac 373
DB 305 CCCCAGACCTCTGTAGACCTGACACATGACACAGCAATCTTAACCTCAAGAAAAAC 364
QY 374 caaacgtaacacaaacccgcccacagagacgtcaagttccggggggtgtgtagtct 433
DB 305 CAACGTAACACCAACCGCCGCCACAGACGTCAAGTTCCGGGCGGTGTGATGCT 424
QY 434 tggtagtcttaccctgttgcgcgcaaggagcccaagttgtgtgcgcgactagaa 493
DB 425 TGGTGAGCTTACTCTGTGCGCGGAGAGGCCCCAGTGTGGTGTGCGGCCGCCAGAA 484
QY 494 gaactccgagcgttcgcaacccgttggaaggcgaacactatcccaaggctgcggcc 553
DB 485 GACTTCGAGCGGTGCGCAACCTGTGAGAGCGACACACCTATCCCAAGGCTCGCGGCC 544
QY 554 cgaaggagagacctgggctcagcccggtatctcttggccctctatggcaagagggcat 613
DB 545 CGAGGCGAGACCTGGGCTAGCCCGGGTACCTTGAGCTCTCTATGGCAATAGGGCTT 604
QY 614 ggggtggcagatggtctcctgtgcgcccggcgtcccgccctagttgggaccttcgga 673
DB 605 AGGTTGGGAGAGAGAGGCTCTGTACACCCCGGCTCCCGGCTAGTTGGGGCCCAAGGA 664
QY 674 ccccgagcgttagtgcgttaatttggtaagtgatcatg 713
DB 665 CCCCAGGCTAGGTCGCGCTAATTGGGTAGGTCATCATGAT 704

RESULT: 7
US-08-384-616-3
Sequence 3, Application US/08384616
Patent No. 5847101
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKE, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
ADDRESS: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616
CLASSIFICATION: 424
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990

RESULT 9
US-09-315-850-3
; Sequence 3, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, HIOTO

```

APPLICANT: Fuke, Isao
APPLICANT: Mori, Chisato
APPLICANT: YAKAMIZAMA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: Armstrong, Westerman, Hattori, Mcleland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,850
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/059,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mcleland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333..1499
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1499
OTHER INFORMATION: /note="sequence = 1 - 1499 of SEQ
OTHER INFORMATION: ID NO: 1"
US-09-315-850-3

```

Matches 689; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 14 tggggcgacactccacatagatcaactcccttgtaggaactactgtcttcaacgagaa 73
 |||||
 Db 5 tggggcgacactccacatagatcaactcccttgtaggaactactgtcttcaacgagaa 64
 |||||

QY 74 agcgtctagcacttggtatagatagtgctgtgcagctccagggccccctcccgag 133
 |||||
 Db 65 agcgtctagcacttggtatagatagtgctgtgcagctccagggccccctcccgag 124
 |||||

QY 134 agagccatagtgctctgcggaacccgttgtaglacccggaattgcagagacagcggtcc 193
 |||||
 Db 125 agagccatagtgctctgcggaacccgttgtaglacccggaattgcagagacagcggtcc 184
 |||||

QY 194 ttcttgatcaaccgcgtcaatgacctgagatlttggtcgtyccccgcgagactgtag 253
 |||||
 Db 185 ttcttgatcaaccgcgtcaatgacctgagatlttggtcgtyccccgcgagactgtag 244
 |||||

QY 254 ccgagtagtggttggtgcggaagcgctgtgtactgtcgtatggtgtgtgtgagtg 313
 |||||
 Db 245 ccgagtagtggttggtgcggaagcgctgtgtactgtcgtatggtgtgtgtgagtg 304
 |||||

QY 314 ccccggaagctctgtagacgtgacacatgacacaaatccctaaacctcaagaaagaaac 373
 |||||
 Db 305 ccccggaagctctgtagacgtgacacatgacacaaatccctaaacctcaagaaagaaac 364
 |||||

QY 374 caacgttaacacacacccgcgcacacaggaacgtcaagltccggggtgtgtcagatgt 433
 |||||
 Db 365 caacgttaacacacacccgcgcacacaggaacgtcaagltccggggtgtgtcagatgt 424
 |||||

QY 434 tggtagatctacgtgtgcgacgagggccccaggtgtgtgtgtgtgtgtgtgtgtgtgt 493
 |||||
 Db 425 tggtagatctacgtgtgcgacgagggccccaggtgtgtgtgtgtgtgtgtgtgtgtgt 484
 |||||

QY 494 gacttcgagcggtctgcacactctgtggaagcgacacacatccccaaggtctgcgagcc 553
 |||||
 Db 485 gacttcgagcggtctgcacactctgtggaagcgacacacatccccaaggtctgcgagcc 544
 |||||

QY 554 cgaagggcaggaactgtgtctcaagcccggtatctctgtgccccctctatggaacagggagat 613
 |||||
 Db 545 cgaagggcaggaactgtgtctcaagcccggtatctctgtgccccctctatggaacagggagat 604
 |||||

QY 614 ggggtgagcaggaatgctctgtgccccgcgaggtcccgagctagttggggcccttcgga 673
 |||||
 Db 605 ggggtgagcaggaatgctctgtgccccgcgaggtcccgagctagttggggcccttcgga 664
 |||||

QY 674 ccccgcgtagtgctgctgtaatttggtlaaggtlcatcgat 713
 |||||
 Db 665 ccccgcgtagtgctgctgtaatttggtlaaggtlcatcgat 704
 |||||

RESULT 10
 US-08-324-977-1
 ; Sequence 1, Application US/08324977
 ; Patent No. 5747339
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKE, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSSEE: Armstrong, Westernman, Hattori, McLealand &
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/324,977
 FILING DATE: 18-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-167466
 FILING DATE: 25-JUN-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-230921
 FILING DATE: 31-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-305605
 FILING DATE: 09-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/039,706
 FILING DATE: 30-JUL-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/769,996
 FILING DATE: 02-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/635,451
 FILING DATE: 28-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Stevens-Smith, Theresa M.
 REGISTRATION NUMBER: 36,281
 REFERENCE/DOCKET NUMBER: 900703D
 TELEPHONE: (202) 659-2930
 TELEFAX: (202) 887-0357
 TELEX: 440142
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9416 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA from genomic RNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 333..9362
 US-08-324-977-1

Query Match 95.7%; Score 682.4; DB 1; Length 9416;
 Best local Similarity 98.4%; Pred. No. 1,9e-197;
 Matches 689; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 14 tggggcgacactccacatagatcaactcccttgtaggaactactgtcttcaacgagaa 73
 |||||
 Db 5 tggggcgacactccacatagatcaactcccttgtaggaactactgtcttcaacgagaa 64
 |||||

QY 74 agcgtctagcacttggtatagatagtgctgtgcagctccagggccccctcccgag 133
 |||||
 Db 65 agcgtctagcacttggtatagatagtgctgtgcagctccagggccccctcccgag 124
 |||||

QY 134 agagccatagtgctctgcggaacccgttgtaglacccggaattgcagagacagcggtcc 193
 |||||
 Db 125 agagccatagtgctctgcggaacccgttgtaglacccggaattgcagagacagcggtcc 184
 |||||

QY 194 ttcttgatcaaccgcgtcaatgacctgagatlttggtcgtyccccgcgagactgtag 253
 |||||
 Db 185 ttcttgatcaaccgcgtcaatgacctgagatlttggtcgtyccccgcgagactgtag 244
 |||||

QY 254 ccgagtagtggttggtgcggaagcgctgtgtactgtcgtatggtgtgtgtgagtg 313
 |||||
 Db 245 ccgagtagtggttggtgcggaagcgctgtgtactgtcgtatggtgtgtgtgagtg 304
 |||||

QY 314 ccccggaagctctgtagacgtgacacatgacacaaatccctaaacctcaagaaagaaac 373
 |||||
 Db 305 ccccggaagctctgtagacgtgacacatgacacaaatccctaaacctcaagaaagaaac 364
 |||||

Sequence 1, Application US/08904686A
Patent No. 5998130
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKE, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 9007036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-0357
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333..9362
US-08-904-686A-1

Query Match 95.7%: Score 682.4; DB 2; Length 9416;
Best Local Similarity 98.4%: Pred. No. 1.9e-192;
Matches 689; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 14 tggggcgacactccacatagatcaccctgtgaggaactactgtcttcacgacagaa 73

|||||
5 TGGGGCGACACTCCACCATAGATCACTCCCTGTGAGCAACTACTGTCTTCACGACGAA 64
QY 74 agcctcagccatgagcgtatagatagtgctgtagagcctcagaacccccctccggg 133
DB 75 ACCGCTACCGCTTGGCGTTAGTATGATGTCTGTGAGCTCCAGGACCCCTCCCGGG 124
QY 134 agagccatagtgctcgcggaacccggtgagtaacccgaattgcagagacgagctcc 193
DB 135 AGAGCATAGTGTCTGCGGAACCGGTGAGTACACGGAAATGCCAGACGACGGGTCC 184
QY 194 ttcttgatcaaccgcctcaatgcctggaagtgtggcggtgcccgagactgtag 253
DB 195 TTTCTTGATCAACCCGCTCAATGCTGAGATTTGGCGGTCCCGCGAGACATGCTAG 244
QY 254 ccgagtagtggtggtgcggaagccctgtgtactgctgataagggtgtgtggaagt 313
DB 255 CCGAGTAGTGTGGGTCCGGAAGCCCTTGTGTACTGCTATAGGCTGCTTGGAGTG 304
QY 314 ccccggaaggtcgtgtagacggtgacacatgagcaaaatcctaactcaagaagaac 373
DB 315 CCCCGGAGGTCTCGTAGACCGTGCACATGAGACAGAACTTAACCTCAAGAAAGAAC 364
QY 374 caacgttaaccacacacgacgacagagacgttaagtcccgagcgtgtgtcaatcgt 433
DB 375 CAACGTAAACACCAACCGCCGACAGACGCTCAAGTTCCCGGCGGTGTGACATGCT 424
QY 434 tggtagagttacgtgtgacgacgagggcccgaggttgggtgctgacgacacagaa 493
DB 435 TGGTAGAGTTTACGTTGTGCCCGCAGGGGCCCCAGTTGGGTGTGCCGCCGCCACGAA 484
QY 494 gacttcgagcgtgtgcgaacctgtggaagcgacaaactatcccaaggtcgcggcc 553
DB 495 GACTTCCGAGCGGTGCAACCTCTGGAAGGAGACAACTATCCCAAGGCTCGCGGCC 544
QY 554 cgaagcgagaccctggtcgaacccgggtatcctgtgccccttatgtgcaagaaggat 613
DB 555 CGAGGCGACAGACTGGCTCAACCCGGGTACCTTGGCTCTCTATGGAAGAGGGCTT 604
QY 614 ggggtgggcaagatlgctctcgtgcccgcgagctcccgacctagtgtgggcccctcga 673
DB 615 AGGTGGCGAGAGTGGCTCTCTCAACCCCGGCGTCCCGGCTATTGGGGCCCGACGGA 664
QY 674 ccccggtgtagtgctggttaatttgggttaaggtatcgtat 713
DB 675 CCCCGGCGTGGTGGCTGCTTAATTTGGTAAGTCAATCAT 704
RESULT 13
US-09-3:5-850-1
Sequence 1, Application US/09315850
Patent No. 6217872
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKE, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0

Query Match 95.1%; Score 678.4; DB 2; Length 1863;
Best Local Similarity 98.4%; Pred. No. 1.5e-191;
Matches 685; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
QY 18 ggcagactccacccatcaatcctccctgtagaagaactcgtcttcaacgagaagcg 77
    |||||||
DB 1 ggcgacactccacccatcaatcctccctgtagaagaactcgtcttcaacgagaagcg 60
QY 78 tctagccatggcgttagatgagtgctgtagcagcctccagagaccccccctccgagag 137
    |||||||
DB 61 tctagccatggcgttagatgagtgctgtagcagcctccagagaccccccctccgagag 120
QY 138 ccatagtgctgtagcagcagtgtagtaacacggaaattgccaagagacccgggtcttctc 197
    |||||||
DB 121 ccatagtgctgtagcagcagtgtagtaacacggaaattgccaagagacccgggtcttctc 180
QY 198 ttgagtaaccccgctcaatgctgtagaattgggggtgcccgcgagaaactgtagccga 257
    |||||||
DB 181 ttgagtaaccccgctcaatgctgtagaattgggggtgcccgcgagaaactgtagccga 240
QY 258 gtagtgctgtagcagaaagccctgtgtactgctgtagaagtgctgtagcagtgccgc 317
    |||||||
DB 241 gtagtgctgtagcagaaagccctgtgtactgctgtagaagtgctgtagcagtgccgc 300
QY 318 gggagtgctgtagaagcgtgtagcagcaaaatcctaaacctcaagaagaaaccaa 377
    |||||||
DB 301 gggagtgctgtagaagcgtgtagcagcaaaatcctaaacctcaagaagaaaccaa 360
QY 378 cgtaacacacaccccgccgacagagcgtgtagaattcccgagcggtgtagaattggt 437
    |||||||
DB 361 cgtaacacacaccccgccgacagagcgtgtagaattcccgagcggtgtagaattggt 420
QY 438 ggaattacactgtgtccgagcagagggcccgagtggtgtgtgtgtgtgtgtgtgtgt 497
    |||||||
DB 421 ggaattacactgtgtccgagcagagggcccgagtggtgtgtgtgtgtgtgtgtgtgt 480
QY 498 tccgagtgctgtagcagcctgtgtggaagcgagacacataccccaagtgctgtagcccgag 557
    |||||||
DB 481 tccgagtgctgtagcagcctgtgtggaagcgagacacataccccaagtgctgtagcccgag 540
QY 558 ggcagagactggcgtcagcccgaggtatcctgtgcccctctatggaagagagcgatgg 617
    |||||||
DB 541 ggcagagactggcgtcagcccgaggtatcctgtgcccctctatggaagagagcgatgg 600
QY 618 tggcagagtgctgtagcagcctgtgtggaagcgagacacataccccaagtgctgtagcccgag 677
    |||||||
DB 601 tggcagagtgctgtagcagcctgtgtggaagcgagacacataccccaagtgctgtagcccgag 660
QY 678 cggcgtagtgctgtagaattggtgtagaagtgatg 713
    |||||||
DB 661 cggcgtagtgctgtagaattggtgtagaagtgatg 696
```

RESULT 15
US-08-988-321B-37
Sequence 37, Application US/08988321B
Patent No. 6174868
GENERAL INFORMATION:
APPLICANT: Kevin P. Anderson et al.
TITLE OF INVENTION: Compositions And Methods For Treatment Of Hepatitis C V
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
City: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM COMPATIBLE

OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORDPERFECT 6.1 FOR WINDOWS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/988, 321B
FILING DATE: December 10, 1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/650, 093
FILING DATE: May 17, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/452, 841
FILING DATE: May 30, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/397, 220
FILING DATE: March 9, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/945, 289
FILING DATE: September 10, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0245
TELECOMMUNICATION INFORMATION:
TELEPHONE: (856) 810-1515
TELEFAX: (856) 810-1454
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 686
TYPE: nucleic acid
STRANDEDNESS: Single
NOPOLOGY: Linear
ANTI-SENSE: No

Query Match 93.1%; Score 663.6; DB 4; Length 686;
Best Local Similarity 80.8%; Pred. No. 2.4e-187;
Matches 554; Conservative 118; Mismatches 14; Indels 0; Gaps 0;

```
QY 1 gcccaccccgatgaggcgacacatccacatagatcactccctgtgaggaactactg 60
    |||||||
DB 1 gccacgccccgaguuuggggggacacatccacatagatcactccctgtgaggaactactg 60
QY 61 tctcaagcagaagaagcgtctgtagcagtggttagatgagtgctgtagcagcctccagag 120
    |||||||
DB 61 ucucacacacagaaagcgucagccagucgucgucgucgucgucgucgucgucgucgucg 120
QY 121 ccccccctccgagagagcagatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
    |||||||
DB 121 ccccccctccgagagagcagatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
QY 181 gacgacgggtgctcttctgtgataacccgctcaatgctgtgagatgtgagtggtgccc 240
    |||||||
DB 181 gacgacgggtgctcttctgtgataacccgctcaatgctgtgagatgtgagtggtgccc 240
QY 241 gacgacgggtgctcttctgtgataacccgctcaatgctgtgagatgtgagtggtgccc 240
    |||||||
DB 241 gacgacgggtgctcttctgtgataacccgctcaatgctgtgagatgtgagtggtgccc 240
QY 301 ggcagactgtagcagcagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtg 300
    |||||||
DB 301 ggcagactgtagcagcagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtg 300
QY 361 ggcagactgtagcagcagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtg 360
    |||||||
DB 361 ggcagactgtagcagcagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtg 360
QY 421 ggcagactgtagcagcagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtg 480
    |||||||
DB 421 ggcagactgtagcagcagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtg 480
QY 481 ggcagactgtagcagcagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtg 540
    |||||||
DB 481 ggcagactgtagcagcagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtg 540
```

```
Db 481 GCCCGAUCAGAGACUUCGAGCGGCCAACCCTGUGAAGCGCAGACCUAUCCCCA 540
QY 541 aggtctgcccggccccgagggcaaggacctgggtctagccgggtatctctggccctctatg 600
    ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 AGGCTUCGGCGCGCCGAGGGCGAGGGCCUGGGCUCAGCCCGGGUAVCCUUGGCCCCCUUAUG 600
QY 601 gcaacgagggcatgggtgtggcagatgtctctgtcgccccggcggtctccggcctagtt 660
    ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 GCADUGAGGGCAUGGGGUGGGCGAGADUGGCUCCUGACCCCGGCUCCCGGCGCUAGUU 660
QY 661 gggggccttcgagcccccgagctagg 686
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 661 GGGGCCCCACAGGACCCCGGCUAGG 686
```

Search completed: March 1, 2002, 15:26:13
Job time: 36631 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 15:24:35 ; Search time 13643.2 Seconds

(without alignments)
413.542 Million cell updates/sec

Title: US-09-763-836-7

Perfect score: 342
1 gccagcccccctgatgggggc.....ggctcctgtagaccctgcacc 342

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: 1: gb_ba: 2: gb_hvg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_om: 20: em_or: 21: em_ov: 22: em_pat: 23: em_ph: 24: em_pl: 25: em_ro: 26: em_sts: 27: em_sy: 28: em_un: 29: em_vl: 30: em_higo_hum: 31: em_higo_inv: 32: em_higo_rod: 33: em_hig_hum: 34: em_hig_inv: 35: em_hig_rod: 36: em_hig_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	342	100.0	9538	14	AB016785	AB016785 Hepatitis
2	335.6	98.1	9596	14	AF054249	AF054249 Hepatitis
3	328.4	96.0	341	14	HPC5UTR	LA0552 Hepatitis C
4	326.8	95.6	2033	6	E08374	E08374 N-terminal
5	326.8	95.6	2033	6	E08869	E08869 5' region of
6	326.8	95.6	4987	6	E08872	E08872 DNA encodin
7	326.8	95.6	9431	14	HPCPCPO	D45172 Hepatitis C
8	326.8	95.6	9431	14	HPCPP	D30613 Hepatitis C
9	326.8	95.6	9471	6	E06261	E06261 CDNA encodi
10	326.8	95.6	9471	6	E06457	E06457 CDNA encodi
11	325.2	95.1	341	6	AR019856	AR019856 Sequence
12	325.2	95.1	9401	6	AR030378	AR030378 Sequence
13	325.2	95.1	9416	6	AR110845	AR110845 Sequence
14	325.2	95.1	9548	14	D89815	D89815 Hepatitis C
15	325.2	95.1	9595	6	AR119832	AR119832 Sequence
16	325.2	95.1	9595	14	AF054247	AF054247 Hepatitis
17	325.2	95.1	9595	14	AF054248	AF054248 Hepatitis
18	325.2	95.1	9599	6	AR119833	AR119833 Sequence
19	323.6	94.6	341	6	AR119855	AR119855 Sequence
20	323.6	94.6	341	6	AR153763	AR153763 Sequence
21	323.6	94.6	341	14	HPC5NR2	D31602 Human Hepat
22	323.6	94.6	341	14	HPCRNAL	M58406 Hepatitis C
23	323.6	94.6	342	6	AR095001	AR095001 Sequence
24	323.6	94.6	349	14	HCY1A5G	Y10150 Hepatitis C
25	323.6	94.6	350	6	AR028132	AR028132 Sequence
26	323.6	94.6	686	6	AR142355	AR142355 Sequence
27	323.6	94.6	780	6	AR095006	AR095006 Sequence
28	323.6	94.6	923	6	AK153761	AK153761 Sequence
29	323.6	94.6	1880	14	HPC5TRJ4	D00832 Hepatitis C
30	323.6	94.6	2540	6	E04260	E04260 CDNA encodi
31	323.6	94.6	2540	6	E04805	E04805 CDNA to 5'
32	323.6	94.6	2540	6	E07391	E07391 CDNA encodi
33	323.6	94.6	7989	6	AX036255	AX036255 Sequence
34	323.6	94.6	7989	12	SSR242652	AJ242652 Hepatitis
35	323.6	94.6	8001	6	AX036254	AX036254 Sequence
36	323.6	94.6	8001	6	AX036257	AX036257 Sequence
37	323.6	94.6	8001	6	AX036259	AX036259 Sequence
38	323.6	94.6	8001	6	AX036261	AX036261 Sequence
39	323.6	94.6	8001	12	SSR242654	AJ242654 Hepatitis
40	323.6	94.6	8637	6	AX036253	AX036253 Sequence
41	323.6	94.6	8637	12	SSR242651	AJ242651 Hepatitis
42	323.6	94.6	8649	6	AX036256	AX036256 Sequence
43	323.6	94.6	8649	12	SSR242653	AJ242653 Hepatitis
44	323.6	94.6	9401	6	E66593	E66593 Hepatitis C
45	323.6	94.6	9401	6	I71894	I71894 Sequence 9

ALIGNMENTS

RESULT: 1
AB016785
LOCUS: AB016785
DEFINITION: Hepatitis C virus genomic RNA, complete sequence.
ACCESSION: AB016785
VERSION: AB016785.1 GI:5821154
KEYWORDS: polypeptide.
SOURCE: Hepatitis C virus
ORGANISM: Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae; Hepacivirus.
REFERENCE: 1 (sites)
AUTHORS: Zhang, J., Yamada, O., Ito, T., Akiyama, M., Hashimoto, Y., Yoshida, H., Makino, R., Masago, A., Uemura, H. and Arai, H.
TITLE: A single nucleotide insertion in the 5'-untranslated region of hepatitis C virus leads to enhanced cap-independent translation
JOURNAL: Virology 261 (2), 263-270 (1999)
MEDLINE: 99428762
REFERENCE: 2 (bases 1 to 9538)
AUTHORS: Zhang, J.

Sept 1, '99

TITLE
JOURNAL

Direct Submission
Submitted (05-AUG-1998) to the DDBJ/EMBL/GenBank databases. Jing
Zhang, Fuso Pharmaceutical Industries, Ltd., Research & Development
Center; 3-30, 2-Chome, Morinomiya, Joto-Ku, Osaka 536-8523, Japan
(E-mail: uemurah@emb.infoweb.ne.jp, Tel: 81-6-969-3131,
Fax: 81-6-964-2706)

FEATURES
source1. 9538
Location/Qualifiers

/organism="Hepatitis C virus"
/db_xref="taxon:11103"
343. .9384

CDS

/codon_start=1
/product="polyprotein"
/protein_id="FA83719.1"
/db_xref="GI:5821155"
/translation="MSTNPKPQKRTKRNTRRPDVKFPGGQIVGVYLLPRRPRRL
GVRAKRTKSEROPGRGROPVPRKARPEGRMAQGPVPLNGEGLGAEMLLSPRG
SRPWSGPTDPRRRSRLNGKVIDTLTGCGADLMGVIPLVAPGAGARALAHGVLEL
GVNATGNLPGCSFSIFLLALSLCTIPASVAVENSGVYHTNDCSSISVGAAD
MIMHPGCVPCVRENSSRCWALTPTLAARRSIPTTIRHVDLLGAAAFCSAM
VGDLCGSPVLSQLFTFSPRRRTVODNCISLPCHVSGHRAMMMMMSPALV
SOLLRIPQAVDMVAGAHGVLAGLAVYSMGNNAKVLVLLFAGVDRTHVVGQ
TGRITLIGTAMFAFGPHQLDLINTNGSWHINRTALNCNDSINTGFLAFAARENS
SGCPERMASCRPIDKFOVQWGPITHAVPDNDLDRPCWHYAPQCGITIPASQVCPVY
CTPSPVYVGTDRGAPVTVNGENETVILLNTRPPOGNGECGTMNNGFAKTCG
GPPCNIGGVNNTLCPTDCPKRHPKATYTCGSPWILPRCMVYPRRLMPCVY
PTIRKRVYVGVVRLTAACMWTGERODLDRSELPLLSITTEMQVPCSTT
LPALSTGLIHQONIVQVYLYGVSAVSVITKVEYILLFLLLADARVCAALMMLL
LQAEEALENLVINAASVAGHGTSEFVEFCANAYIKGRVPAAGAAFGVCPPL
LLVLLPRAVAGREMAASCGAVLGLVLLTSPHKLPRLLIMLYETIRAEAL
LVQVYIPPLNVRGSDAILLTCVAPHGIFEVTLTLLAIGTLLIAGLIRVYFV
RAOGLIRACMLVRKVAAGHYVOMAFKMLAALGTGYVHTHPLDOMAHGLDLAVY
EYVSDMETKITITWCAPTAACGDIISCIPIASARGRELLGAPASFERGRLLAPI
TAYSOOTGLASITSLTGRDNKRVBEGVOVATSPATSCGSDLVYVTRHADVTPRR
KTLAPGKPIQMTYTNVDODLVGWPAPSGARSLCTCGSSDLVYVTRHADVTPRR
GDSRGLSPRPVSYLKSGSGGLPCSGHGVGIFRAVCTGVAKAAVDFESEMET
TMRSEVFTDNSSPAPVOTFOYALHAPTSKSTKVAAYAGKVLVLPVSAAT
LGFAYMSKANGIDINIRGVRTITTGAPITVSTYGRFLADGCGGAYDITIDCEH
STDSITIGTGVLDQETAGARLVLAATPDSGVTPHPIEVALSNCEIIFPG
KAIPETIKGCHHLIFCSKSKCDLAKSLKSLGNAVYRGLDVSVIPSGDVAV
AIDALMTGYTDFSDVTCNFCVOTVPSLDPTPTTLETTHVPODSVRSORGRHTR
GRGGIYRFVIGERSPGMSDFVCECYDAGAWELPAETVLAAYLNPGLPFG
ODHLEFMEGVFTGLTHIDAHFISOTKQAGDNPEYVAYQATVCAAOAPPSMDQMK
CLIRKPTLHGPPTLLYRLGAVONETLTHPTTKYIKTMSADLEVYVSTWYVGLV
AALAYCLITGTSVYVGRILLISGRPAIIPDEPVYIREDEMECSHPIYEGQILA
EOKKALGILLQATKQAAAPVYESKWRLEAWAHMMEISGIQYLAGLSTPLG
NPALSLMAFTASITSPITQHTLTLNGLGVAOAPPSAASAFVAGALVAGAVS
IGVAKVLVDLAGVAGVAGALVAFKVNSEVPSTEDLVNLLALISGALVAGVCA
AILRHVPGEGAVOMVNRILAFASRGNHVPTHEVPESDAAARTQIISLITLQIL
KILHWINEDSCPSGSMLRDVMICTVLEDFETWOSLALRPLPCVPTLSCORGY
RGVNRDGIIMHTPCGCAOITGVHNGSMRIYVPTCSNMTHGTPITNTYTTGCTPS
PAPVSKALMVAEYEVYRVGDVPHVTCMTTDNVACPCOVAPAEFFTELQVRLH
RYAPACKFLRDEVTLQVGLNOYVGSOLPCEPEPVDVILSMIDPSHITAEAKRR
LARGSVSPSLASSASQSLASLAKATCTTHSDADALIEANILMROEMGNITRVE
SENVRVILDSFDPLAEDEREISVAADILRKTKEFPAMPIMARPDYNPPLSEMN
PDYVPVHVGCPISPTRAPPIPPRRKRTVLTESAVSALAEATKTFGSESVAVC
SGTAPAPGSSDDGCTGSDEYESSMPLRGEOPDPLSDGSMSTVESESESVCC
SMSTYTCALITPCAARESKIPINALSILRHHNVYATTSRASQOKVATEERLO
VLDDHYRVLEEMAKASTVAKLLSVEACKLTPHSHARKFGAGADVNLSEKAV
NLHISVMDLLEDEPTPIINTIMAKNEVCOPEPGRPARLIVPDLGVAVKEMA
LYDAVSTLPQAVNGSSYFOYSPGQVFEVLNAMSKNPMGFAVDAFCSTVTEEN
LIVREESIYOCCLAPPEAROIRSLTERLYIGGLPNSGONGCAGRCASVITTEEN
NTLCHLKASACAAKACDCTMLVCGDILVICESACTORDASLARTFAMTRISA
PPGDPPOPEYELITGSSSVSAHDSGRVYITLTDPTPLARAKMETARTHPVN
SWLGNITMTATLWARMLMTHFSLIAQOSLEALDCCQYGAATYSTEPDLQILIO
RLHGLSFLSHSYSGEINVAQCLRKGVPLRWRHARSVRKRLISOGGRATCG
KYLEFNVAVRTKLKLTPIPAASQDLDSGWFAVYSGGDIYHSLSRARPRFMWCLLIS
VGVGILYLPNR"

BASE COUNT 1908 a 2882 c 2723 g 2025 t
ORIGIN

Query Match

100.0%; score 342; DB 14; length 9538;

Best local Similarity 100.0%; Pred. No. 1.8e-86;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gcaagcccccgtgtgggggacacccaccacccatagatcccccgtgggaactac 60
Db 1 GCCAGCCCGCTGATGGGGGGACACTCCACCATAGATCACTCCCTGTGAGACTACTG 60
QY 61 lcttcaagcaagaagcgtctagccatagcglltagtagtgcgtgcagccctcagcc 120
Db 61 TCTTACAGCGAAGACGCTAGACCATAGCGCTTAGTATGATGTGTCAGACCTCCAGGCC 120
QY 121 cccccctcccggaagagccatagtgctgcggaaaccgtgagtaacacggaaattccag 180
Db 121 CCCCCCTCCCGGAGAGCATAGTGTGTCGGAGACCGGTGAGTACACCGGAATTCCAG 180
QY 131 gaagcagcggctctctcttgatcaatcccgcccaatgctctgagattggcgctgcacc 240
Db 131 GAGACCGGGGTCTCTTCTTGATCAATCCCGCTCAATGCTGAGATTGGGCTGCCCC 240
QY 241 cgcgaagactgtagccgagtagtgcgtgcgaaagcccttgtagctagcctatag 300
Db 241 CGCGAGACCTGCTTAGCGGAGTAGTGTGGTCCGCAAGGCCCTTGCTAGCTGATGAG 300
QY 301 ggtgcttgagagtgcccgaggagctcgtgaagccgtgacc 342
Db 301 GGTGCTTGCGAGTGCCCCGGAGGCTCTGTAGCGCTGACC 342

RESULT 2

LOCUS AF054249 9596 bp RNA VRL 02-MAY-1998
DEFINITION Hepatitis C virus HC-J4, pcv-J4L4S, complete genome.
ACCESSION AF054249
VERSION AF054249.1 GI:3098636
KEYWORDS
SOURCE
ORGANISM Hepatitis C virus.
Hepatitis C virus.

REFERENCE 1 (bases 1 to 9596)
AUTHORS Yanagi, M., St. Claire, M., Shapiro, M., Emerson, S. U., Purcell, R. H. and
Bukh, J.
TITLE Transcripts of a chimeric CDNA clone of hepatitis C virus genotype
1b are infectious in vivo
JOURNAL Virology 244 (1), 161-172 (1998)
MEDLINE 98240944
REFERENCE 2 (bases 1 to 9596)
AUTHORS Yanagi, M. and Bukh, J.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-1998) Hepatitis Viruses Section, LID, NIAID, NIH,
Center Dr., Building 7, Room 201, Bethesda, MD 20892, USA
FEATURES
source
location/Qualifiers
1. .9596
/organism="Hepatitis C virus"
/strain="HC-J4"
/isolate="pcv-J4L4S"
/db_xref="taxon:11103"
/note="non-infectious full-length cDNA"
343. .9375
/codon_start=1
/product="polyprotein"
/protein_id="FA83724.1"
/db_xref="GI:3098637"
/translation="MSTNPKPQKRTKRNTRRPDVKFPGGQIVGVYLLPRRPRRL
GVRAKRTKSEROPGRGROPVPRKARPEGRMAQGPVPLNGEGLGAEMLLSPRG
SRPWSGPTDPRRRSRLNGKVIDTLTGCGADLMGVIPLVAPGAGARALAHGVLEL
GVNATGNLPGCSFSIFLLALSLCTIPASVAVENSGVYHTNDCSSISVGAAD
MIMHPGCVPCVRENSSRCWALTPTLAADASVPTTIRHVDLLGTAACFSAM
VGDLCGSPVLSQLFTFSPRRRTVODNCISLPCHVSGHRAMMMMMSPALV
SOLLRIPQAVDMVAGAHGVLAGLAVYSMGNNAKVLVLLFAGVDRTHVVGQ
TGRITLIGTAMFAFGPHQLDLINTNGSWHINRTALNCNDSINTGFLAFAARENS
SGCPERMASCRPIDKFOVQWGPITHAVPDNDLDRPCWHYAPQCGITIPASQVCPVY
CTPSPVYVGTDRGAPVTVNGENETVILLNTRPPOGNGECGTMNNGFAKTCG
GPPCNIGGVNNTLCPTDCPKRHPKATYTCGSPWILPRCMVYPRRLMPCVY
PTIRKRVYVGVVRLTAACMWTGERODLDRSELPLLSITTEMQVPCSTT
LPALSTGLIHQONIVQVYLYGVSAVSVITKVEYILLFLLLADARVCAALMMLL
LQAEEALENLVINAASVAGHGTSEFVEFCANAYIKGRVPAAGAAFGVCPPL
LLVLLPRAVAGREMAASCGAVLGLVLLTSPHKLPRLLIMLYETIRAEAL
LVQVYIPPLNVRGSDAILLTCVAPHGIFEVTLTLLAIGTLLIAGLIRVYFV
RAOGLIRACMLVRKVAAGHYVOMAFKMLAALGTGYVHTHPLDOMAHGLDLAVY
EYVSDMETKITITWCAPTAACGDIISCIPIASARGRELLGAPASFERGRLLAPI
TAYSOOTGLASITSLTGRDNKRVBEGVOVATSPATSCGSDLVYVTRHADVTPRR
KTLAPGKPIQMTYTNVDODLVGWPAPSGARSLCTCGSSDLVYVTRHADVTPRR
GDSRGLSPRPVSYLKSGSGGLPCSGHGVGIFRAVCTGVAKAAVDFESEMET
TMRSEVFTDNSSPAPVOTFOYALHAPTSKSTKVAAYAGKVLVLPVSAAT
LGFAYMSKANGIDINIRGVRTITTGAPITVSTYGRFLADGCGGAYDITIDCEH
STDSITIGTGVLDQETAGARLVLAATPDSGVTPHPIEVALSNCEIIFPG
KAIPETIKGCHHLIFCSKSKCDLAKSLKSLGNAVYRGLDVSVIPSGDVAV
AIDALMTGYTDFSDVTCNFCVOTVPSLDPTPTTLETTHVPODSVRSORGRHTR
GRGGIYRFVIGERSPGMSDFVCECYDAGAWELPAETVLAAYLNPGLPFG
ODHLEFMEGVFTGLTHIDAHFISOTKQAGDNPEYVAYQATVCAAOAPPSMDQMK
CLIRKPTLHGPPTLLYRLGAVONETLTHPTTKYIKTMSADLEVYVSTWYVGLV
AALAYCLITGTSVYVGRILLISGRPAIIPDEPVYIREDEMECSHPIYEGQILA
EOKKALGILLQATKQAAAPVYESKWRLEAWAHMMEISGIQYLAGLSTPLG
NPALSLMAFTASITSPITQHTLTLNGLGVAOAPPSAASAFVAGALVAGAVS
IGVAKVLVDLAGVAGVAGALVAFKVNSEVPSTEDLVNLLALISGALVAGVCA
AILRHVPGEGAVOMVNRILAFASRGNHVPTHEVPESDAAARTQIISLITLQIL
KILHWINEDSCPSGSMLRDVMICTVLEDFETWOSLALRPLPCVPTLSCORGY
RGVNRDGIIMHTPCGCAOITGVHNGSMRIYVPTCSNMTHGTPITNTYTTGCTPS
PAPVSKALMVAEYEVYRVGDVPHVTCMTTDNVACPCOVAPAEFFTELQVRLH
RYAPACKFLRDEVTLQVGLNOYVGSOLPCEPEPVDVILSMIDPSHITAEAKRR
LARGSVSPSLASSASQSLASLAKATCTTHSDADALIEANILMROEMGNITRVE
SENVRVILDSFDPLAEDEREISVAADILRKTKEFPAMPIMARPDYNPPLSEMN
PDYVPVHVGCPISPTRAPPIPPRRKRTVLTESAVSALAEATKTFGSESVAVC
SGTAPAPGSSDDGCTGSDEYESSMPLRGEOPDPLSDGSMSTVESESESVCC
SMSTYTCALITPCAARESKIPINALSILRHHNVYATTSRASQOKVATEERLO
VLDDHYRVLEEMAKASTVAKLLSVEACKLTPHSHARKFGAGADVNLSEKAV
NLHISVMDLLEDEPTPIINTIMAKNEVCOPEPGRPARLIVPDLGVAVKEMA
LYDAVSTLPQAVNGSSYFOYSPGQVFEVLNAMSKNPMGFAVDAFCSTVTEEN
LIVREESIYOCCLAPPEAROIRSLTERLYIGGLPNSGONGCAGRCASVITTEEN
NTLCHLKASACAAKACDCTMLVCGDILVICESACTORDASLARTFAMTRISA
PPGDPPOPEYELITGSSSVSAHDSGRVYITLTDPTPLARAKMETARTHPVN
SWLGNITMTATLWARMLMTHFSLIAQOSLEALDCCQYGAATYSTEPDLQILIO
RLHGLSFLSHSYSGEINVAQCLRKGVPLRWRHARSVRKRLISOGGRATCG
KYLEFNVAVRTKLKLTPIPAASQDLDSGWFAVYSGGDIYHSLSRARPRFMWCLLIS
VGVGILYLPNR"

CDS

/organism="Hepatitis C virus"
/strain="HC-J4"
/isolate="pcv-J4L4S"
/db_xref="taxon:11103"
/note="non-infectious full-length cDNA"
343. .9375
/codon_start=1
/product="polyprotein"
/protein_id="FA83724.1"
/db_xref="GI:3098637"
/translation="MSTNPKPQKRTKRNTRRPDVKFPGGQIVGVYLLPRRPRRL
GVRAKRTKSEROPGRGROPVPRKARPEGRMAQGPVPLNGEGLGAEMLLSPRG
SRPWSGPTDPRRRSRLNGKVIDTLTGCGADLMGVIPLVAPGAGARALAHGVLEL
GVNATGNLPGCSFSIFLLALSLCTIPASVAVENSGVYHTNDCSSISVGAAD
MIMHPGCVPCVRENSSRCWALTPTLAADASVPTTIRHVDLLGTAACFSAM
VGDLCGSPVLSQLFTFSPRRRTVODNCISLPCHVSGHRAMMMMMSPALV
SOLLRIPQAVDMVAGAHGVLAGLAVYSMGNNAKVLVLLFAGVDRTHVVGQ
TGRITLIGTAMFAFGPHQLDLINTNGSWHINRTALNCNDSINTGFLAFAARENS
SGCPERMASCRPIDKFOVQWGPITHAVPDNDLDRPCWHYAPQCGITIPASQVCPVY
CTPSPVYVGTDRGAPVTVNGENETVILLNTRPPOGNGECGTMNNGFAKTCG
GPPCNIGGVNNTLCPTDCPKRHPKATYTCGSPWILPRCMVYPRRLMPCVY
PTIRKRVYVGVVRLTAACMWTGERODLDRSELPLLSITTEMQVPCSTT
LPALSTGLIHQONIVQVYLYGVSAVSVITKVEYILLFLLLADARVCAALMMLL
LQAEEALENLVINAASVAGHGTSEFVEFCANAYIKGRVPAAGAAFGVCPPL
LLVLLPRAVAGREMAASCGAVLGLVLLTSPHKLPRLLIMLYETIRAEAL
LVQVYIPPLNVRGSDAILLTCVAPHGIFEVTLTLLAIGTLLIAGLIRVYFV
RAOGLIRACMLVRKVAAGHYVOMAFKMLAALGTGYVHTHPLDOMAHGLDLAVY
EYVSDMETKITITWCAPTAACGDIISCIPIASARGRELLGAPASFERGRLLAPI
TAYSOOTGLASITSLTGRDNKRVBEGVOVATSPATSCGSDLVYVTRHADVTPRR
KTLAPGKPIQMTYTNVDODLVGWPAPSGARSLCTCGSSDLVYVTRHADVTPRR
GDSRGLSPRPVSYLKSGSGGLPCSGHGVGIFRAVCTGVAKAAVDFESEMET
TMRSEVFTDNSSPAPVOTFOYALHAPTSKSTKVAAYAGKVLVLPVSAAT
LGFAYMSKANGIDINIRGVRTITTGAPITVSTYGRFLADGCGGAYDITIDCEH
STDSITIGTGVLDQETAGARLVLAATPDSGVTPHPIEVALSNCEIIFPG
KAIPETIKGCHHLIFCSKSKCDLAKSLKSLGNAVYRGLDVSVIPSGDVAV
AIDALMTGYTDFSDVTCNFCVOTVPSLDPTPTTLETTHVPODSVRSORGRHTR
GRGGIYRFVIGERSPGMSDFVCECYDAGAWELPAETVLAAYLNPGLPFG
ODHLEFMEGVFTGLTHIDAHFISOTKQAGDNPEYVAYQATVCAAOAPPSMDQMK
CLIRKPTLHGPPTLLYRLGAVONETLTHPTTKYIKTMSADLEVYVSTWYVGLV
AALAYCLITGTSVYVGRILLISGRPAIIPDEPVYIREDEMECSHPIYEGQILA
EOKKALGILLQATKQAAAPVYESKWRLEAWAHMMEISGIQYLAGLSTPLG
NPALSLMAFTASITSPITQHTLTLNGLGVAOAPPSAASAFVAGALVAGAVS
IGVAKVLVDLAGVAGVAGALVAFKVNSEVPSTEDLVNLLALISGALVAGVCA
AILRHVPGEGAVOMVNRILAFASRGNHVPTHEVPESDAAARTQIISLITLQIL
KILHWINEDSCPSGSMLRDVMICTVLEDFETWOSLALRPLPCVPTLSCORGY
RGVNRDGIIMHTPCGCAOITGVHNGSMRIYVPTCSNMTHGTPITNTYTTGCTPS
PAPVSKALMVAEYEVYRVGDVPHVTCMTTDNVACPCOVAPAEFFTELQVRLH
RYAPACKFLRDEVTLQVGLNOYVGSOLPCEPEPVDVILSMIDPSHITAEAKRR
LARGSVSPSLASSASQSLASLAKATCTTHSDADALIEANILMROEMGNITRVE
SENVRVILDSFDPLAEDEREISVAADILRKTKEFPAMPIMARPDYNPPLSEMN
PDYVPVHVGCPISPTRAPPIPPRRKRTVLTESAVSALAEATKTFGSESVAVC
SGTAPAPGSSDDGCTGSDEYESSMPLRGEOPDPLSDGSMSTVESESESVCC
SMSTYTCALITPCAARESKIPINALSILRHHNVYATTSRASQOKVATEERLO
VLDDHYRVLEEMAKASTVAKLLSVEACKLTPHSHARKFGAGADVNLSEKAV
NLHISVMDLLEDEPTPIINTIMAKNEVCOPEPGRPARLIVPDLGVAVKEMA
LYDAVSTLPQAVNGSSYFOYSPGQVFEVLNAMSKNPMGFAVDAFCSTVTEEN
LIVREESIYOCCLAPPEAROIRSLTERLYIGGLPNSGONGCAGRCASVITTEEN
NTLCHLKASACAAKACDCTMLVCGDILVICESACTORDASLARTFAMTRISA
PPGDPPOPEYELITGSSSVSAHDSGRVYITLTDPTPLARAKMETARTHPVN
SWLGNITMTATLWARMLMTHFSLIAQOSLEALDCCQYGAATYSTEPDLQILIO
RLHGLSFLSHSYSGEINVAQCLRKGVPLRWRHARSVRKRLISOGGRATCG
KYLEFNVAVRTKLKLTPIPAASQDLDSGWFAVYSGGDIYHSLSRARPRFMWCLLIS
VGVGILYLPNR"

PCNIGVGNFTLICPTDCFRKHBEATYTKGSSGMLIPCLVDPYLYLMMYPCILNF
 STEKRVTVGVGEHRLNACWMTGERCNLEDRDSELSPLSTTEMOILPCATITL
 PALSTGLIHLHONIVDOYLXGVGSFVFAIKMEYLLEFLLLADARVACALMMMLL
 LIAOEALLENLVNAASAGAHGILSLFVEPCAMVTKGRILAPGAAYFVGMPLL
 LLALPRAVALDREMAASCGAVLGVFLPLISYVYVRLIMLYOETREBAH
 MOWWPLNVRGDRAILLTCVHPELITITKLLAILLPIWLOQGITRVPVFR
 AOGILRACMLVRKAGHYVOMAFKLCALITGTYNNHILTRDMHNGDNLAAVE
 PVPFSAMETKVTWAGDIAACGDILGLPVARKELEFLPADLSEQGMLLAPIT
 AVSOOTRGVLCVITSLTRDRKNOVEGVQVATQSFATCINGVMTYVHGSGK
 TLAGKPIQMTYNTVDLIVGMOAPPARMTPCSCSSDLYLVTRADVIVRRG
 DSRGSLSPRVSYLKSSGGFLCPSGHVGVFAATGVAACVAVIPESMETIT
 MRSPTDNSSPRYALPQPPFOVHLHAHPGSKSTVPAVAAQKAVIYINVSATL
 GFGATMSKAGHIDENIRTVRTITGGSTITSTYKFLADGSSGAGADITICDCSH
 TDSITLIGIVLDEHETAGALVLAATPVGSTVHPNIEITLGNNGEIKAPIT
 AIPLEAIGRHLJPCHSKRCODELAALVGLAVALYKRLDVSVPIDVYVVA
 TDLMTGTGDFDSVIDCNTCTVOTVDSLDPTFIETTVQDVAZSORGRGRG
 RSGIARTVTPGPRSGMFDSSVLCGYAGACMHLETPAETSRLRAYLNPGLPVC
 DHPFESVPTGLTHIDHFLSOTKQADNFPYIYAQNTCARQAPPSMDQMKC
 LIRKPTILGVTPLIRIGAVONEVITLHPITKYIMACHSADLEVITVWLVGVLA
 ALAACLTGTVVIVGRILISGKPAVVDREVLXOEFDEMECASQLYIEGQMLAE
 OFKQALGLQATKOAEAAAPVSESKRALETPWAKHMMNEISCIQYIAGISTLPGN
 PAISAMAFITASITSPLETONLIFENIIGWAAQALAPPSAASAFVAGIAGAAGSI
 GLKVLVDILAGYGAVGALVAFKVMGSEVPSTEDVNLIPALISPGALVGVCAA
 IIRHVGBEGANOMNRLIAPASGNHVSPPHYBESDAARVQIIPSLITOLIK
 RLHQRINDECSPTCGSWLDRVMDICTVLDKFWDSKLRLEPGVFTSCQNGIK
 GVMRGDGMOTTCPCGAQIAGHVKNGSMRIVGPTCSWMTGTEPINATYTGCPSP
 APNSRALMRVAEEVYEVTRGDVPHVYGTQVTDVWKCPCVQPAEPFEDGVGLHR
 YAPACRPLREDVFOVGINOYLIVGSOJLCEPEPVYUHTSLMDRSHITETARRL
 ARGSPPLASSASOLASPLKATCTHSDDALILANILMROEMGINTETRVSEN
 KVLIDSEFPELHAEGDEREIVAAELIKSRKRPALITMARPDNPLLEESMRODPT
 VPPYHGCPLPPTKAPPIPPRRKRTVYLTESNSALAEATKTFEGSSAVDSGT
 ATALPDLASDDGSDVESHSMPLEGEPPDPLSGSMVSEASEDVCCSMS
 YTNWTCALLTPCAAEESKLPIINLSNLRHNNHVAATYRSLSQKQVIEDRLVLD
 DHYRVLKEMKAKASTYKAKLISIEACKLIPPHAKSGFCGADVNLSRAVNI
 RSVWDLLEDETETPIDTITMAKSEVFCVQPEGGKCPMLIEPGLVGYVCESMLY
 VSTPLQAVMGSVGFQYSPKQVREFLVNTKSKCPKPEVSTDEPSTQTESIRY
 EESTIQCDDLAPBAROAIRSLTERLYTIGPILNKGQNCGRTRCASGVLTTCGNTL
 TCYELATACRAKALQDCTMLVNGDVLVIESATQIDDAALRTFAMRTYAPG
 DPPOEYDELITSCSSNVSAHADSGRYVLLTPDPTPLARAMEFARHPINSMI
 GNIMYAPTLMARMLIMHMFISILAOEOLACOKFLGACQILIEDLDPQIIEHL
 GLSAFTLHSPGSEINRVASCRKLGVPLRFMRBARASVARKLISLOGGRATGCRYL
 FMMAVETKILPIPIPAASOLDISGFVAGVSGDLYHLSLRARPMFPLCLLLESVG
 GYLLPNR"

BASE COUNT 1927 a 2852 c 2706 g 2111 t
 ORIGIN

Query Match 98.1%; Score 335.6; DB 14; Length 9596;
 Best Local Similarity 98.8%; Pred. No. 1.1e-84;
 Matches 338; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gccagcccccctgtagggcgacacccacacatagactcccccgttagaagaactactg 60
 DB 1 gccagcccccctgtagggcgacacccacacatagactcccccgttagaagaactactg 60
 QY 61 tctcaacgagaagaagcgtctagccatgagcgttagtagtgcgtgcaagcctccagcc 120
 DB 61 tctcaacgagaagaagcgtctagccatgagcgttagtagtgcgtgcaagcctccagcc 120
 QY 61 tctcaacgagaagaagcgtctagccatgagcgttagtagtgcgtgcaagcctccagcc 120
 DB 61 tctcaacgagaagaagcgtctagccatgagcgttagtagtgcgtgcaagcctccagcc 120
 QY 121 cccccctcccgagagagccatagtgctgtaggaacggtagtagtaacacggaaattgccag 180
 DB 121 cccccctcccgagagagccatagtgctgtaggaacggtagtagtaacacggaaattgccag 180
 QY 121 cccccctcccgagagagccatagtgctgtaggaacggtagtagtaacacggaaattgccag 180
 DB 121 cccccctcccgagagagccatagtgctgtaggaacggtagtagtaacacggaaattgccag 180
 QY 181 gacgacggggtccttctctggaatcccgctcaatgctgtagaagtctggagcgttcccc 240
 DB 181 gacgacggggtccttctctggaatcccgctcaatgctgtagaagtctggagcgttcccc 240
 QY 241 cgcgagagctgtagccagtagtggtgcggaagaagccttgtagtactgcctgtag 300
 DB 241 cgcgagagctgtagccagtagtggtgcggaagaagccttgtagtactgcctgtag 300
 QY 301 ggtgctgtagtgtagcccgaggaggtctgtagacgtagcacc 342
 DB 301 ggtgctgtagtgtagcccgaggaggtctgtagacgtagcacc 342

DB 301 GGTGCTTCCGAGTCCCGCGGAGGTCTGCTAGACCGTGACCC 342
 RESULT 3
 LOCUS HPC5UTR 341 bp ss-RNA VRL 11-MAR-1995
 DEFINITION Hepatitis C virus gene, 5' UTR.
 ACCESSION L40552
 VERSION L40552.1 GI:705350
 SOURCE
 ORGANISM Hepatitis C virus (strain Taiwanese) cDNA to genomic RNA.
 VIRUSES: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepaciviruses.
 REFERENCE 1 (bases 1 to 341)
 YEN, J.-H., CHANG, S.-C., HU, C.-R., CHU, S.-C., LIN, S.-S., HALEH, Y.-S.,
 AND CHANG, M.-P. Cellular proteins specifically bind to the 5'-noncoding region of
 hepatitis C virus RNA
 JOURNAL Virology (1995) In press
 FEATURES
 source 1. 341
 protein_bind /organism="Hepatitis C virus"
 5'UTR /strain="Taiwanese"
 1. 341 /db_xref="taxon:11103"
 protein_bind /bound_moiety="p120 and p87"
 131. 253 /db_xref="taxon:11103"
 103 c 106 g 69 t
 BASE COUNT 63 a 103 c 106 g 69 t
 ORIGIN
 Query Match 96.0%; Score 328.4; DB 14; Length 341;
 Best Local Similarity 99.4%; Pred. No. 2.1e-82;
 Matches 340; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 gccagcccccctgtagggcgacacccacacatagactcccccgttagaagaactactg 60
 DB 1 gccagcccccctgtagggcgacacccacacatagactcccccgttagaagaactactg 60
 QY 61 tctcaacgagaagaagcgtctagccatgagcgttagtagtgcgtgcaagcctccagcc 120
 DB 61 tctcaacgagaagaagcgtctagccatgagcgttagtagtgcgtgcaagcctccagcc 120
 QY 61 tctcaacgagaagaagcgtctagccatgagcgttagtagtgcgtgcaagcctccagcc 120
 DB 61 tctcaacgagaagaagcgtctagccatgagcgttagtagtgcgtgcaagcctccagcc 120
 QY 121 cccccctcccgagagagccatagtgctgtaggaacggtagtagtaacacggaaattgccag 180
 DB 121 cccccctcccgagagagccatagtgctgtaggaacggtagtagtaacacggaaattgccag 180
 QY 121 cccccctcccgagagagccatagtgctgtaggaacggtagtagtaacacggaaattgccag 180
 DB 121 cccccctcccgagagagccatagtgctgtaggaacggtagtagtaacacggaaattgccag 180
 QY 181 gacgacggggtccttctctggaatcccgctcaatgctgtagaagtctggagcgttcccc 240
 DB 181 gacgacggggtccttctctggaatcccgctcaatgctgtagaagtctggagcgttcccc 240
 QY 241 cgcgagagctgtagccagtagtggtgcggaagaagccttgtagtactgcctgtag 300
 DB 241 cgcgagagctgtagccagtagtggtgcggaagaagccttgtagtactgcctgtag 300
 QY 301 ggtgctgtagtgtagcccgaggaggtctgtagacgtagcacc 342
 DB 301 ggtgctgtagtgtagcccgaggaggtctgtagacgtagcacc 342
 RESULT 4
 LOCUS E08374
 DEFINITION N-terminal region of HCV gene.
 ACCESSION E08374
 VERSION E08374.1 GI:2176491
 KEYWORDS JP 199431185-A/1.
 SOURCE Hepatitis C virus.
 ORGANISM Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepaciviruses.
 REFERENCE 1 (bases 1 to 2033)

AUTHORS Seki, M., Honda, Y. and Yamada, S.
 TITLE ANTI-SENSE COMPOUND COMPLEMENTARY WITH GENE OF HEPATITIS C VIRUS
 JOURNAL Patent: JP 1994311885-A 1 08-NOV-1994;
 MITSUBISHI KASEI CORP
 COMMENT OS Unknown (hepatitis C virus)
 PN JP 1994311885-A/1
 PD 08-NOV-1994
 PR 09-AUG-1993 JP 1993217095
 PE 25-AUG-1992 JP 92P 248796, 03-MAR-1993 JP 93P 42736 PI
 SEKI MAKOTO, HONDA YOSHIKAZU, YAMADA SUGURU
 PC C12N15/51, A61K39/29, C12N15/11, C12N15/86;
 CC strandedness: Double;
 CC topology: Linear;
 FH key Location/Qualifiers

FT source 1. .2033
 FT /organism='Unclassified'
 FT complement(107. .199)
 FT /note='antisense DNA fragment containing 10 to 34
 FT nucleotides of this region inhibits FT
 FT translation of HCV
 FT gene'
 FT complement(250. .401)
 FT /note='antisense DNA fragment containing 10 to 34
 FT nucleotides of this region inhibits FT
 FT translation of HCV
 FT gene'
 FT complement(808. .859)
 FT /note='antisense DNA fragment containing 10 to 34
 FT nucleotides of this region inhibits FT
 FT translation of HCV
 FT gene'
 FT Location/Qualifiers
 FT 1. .2033
 FT /organism='Hepatitis C virus'
 FT /db_xref='taxon:11103'
 FT BASE COUNT 379 a 612 c 595 g 447 t
 FT ORIGIN

Query Match 95.6%; Score 326.8; DB 6; Length 2033;
 Best Local Similarity 99.1%; Pred. No. 4.5e-82;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 1 gccagccccctgatgaggcgacactccacatagatcactccctgtgaggaactactg 60
 DB 28 gccagccccctgatgaggcgacactccacatagatcactccctgtgaggaactactg 87
 QY 61 tcttcacgcaaaagcgctctagccatgagcgtttagatagtgctgtgagcctccagcc 120
 DB 88 tcttcacgcaaaagcgctctagccatgagcgtttagatagtgctgtgagcctccagcc 147
 QY 121 cccccctccggagagacatagtgctgtcggaacgggtagtaacacggaattgcccag 180
 DB 148 cccccctccggagagacatagtgctgtcggaacgggtagtaacacggaattgcccag 207
 QY 181 gacgacgggtctcttcttgatcaatccgcctcaatgctcggaatttggcgctgccc 240
 DB 208 gacgacgggtctcttcttgatcaatccgcctcaatgctcggaatttggcgctgccc 266
 QY 241 cgcgagactgctagccgagtagtgctgtcggaacgggttagtaacacggaattgcccag 300
 DB 267 cgcgagactgctagccgagtagtgctgtcggaacgggttagtaacacggaattgcccag 326
 QY 301 ggtctgtgagagtgccccggagaggtctcgtagaccgtgacac 342
 DB 327 ggtctgtgagagtgccccggagaggtctcgtagaccgtgacac 368
 RESULT 5

E08869
 LOCUS E08869 2033 bp RNA
 DEFINITION 5' region of HCV genome.
 ACCSSION E08869
 VERSION E08869.1 GI:2176973
 KEYWORDS JP 1995069899-A/1.
 SOURCE Hepatitis C virus.
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus.

REFERENCE 1 (bases 1 to 2033)
 AUTHORS Seki, M., Honda, Y. and Yamada, S.
 TITLE ANTI-SENSE COMPOUND COMPLEMENTARY WITH GENE OF HEPATITIS C VIRUS
 JOURNAL Patent: JP 1995069899-A 1 14-MAR-1995;
 MITSUBISHI KASEI CORP

COMMENT OS Unknown (Hepatitis C virus)
 PN JP 1995069899-A/1
 PD 14-MAR-1995
 PE 02-SEP-1993 JP 1993241973
 PI SEKI MAKOTO, HONDA YOSHIKAZU, YAMADA SUGURU
 PC A61K31/70, A61K48/00, C07K7/00, C12N15/09//C07H21/04; CC
 strandedness: Double;
 CC topology: Linear;
 FH key Location/Qualifiers

FT source 1. .2033
 FT /organism='Unclassified'
 FT /clone='77N1-19'
 FT /product='N-terminal region of HCV protein'.
 FT Location/Qualifiers
 FT 1. .2033
 FT /organism='Hepatitis C virus'
 FT /db_xref='taxon:11103'
 FT BASE COUNT 379 a 612 c 595 g 447 t
 FT ORIGIN

Query Match 95.6%; Score 326.8; DB 6; Length 2033;
 Best Local Similarity 99.1%; Pred. No. 4.5e-82;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 1 gccagccccctgatgaggcgacactccacatagatcactccctgtgaggaactactg 60
 DB 28 gccagccccctgatgaggcgacactccacatagatcactccctgtgaggaactactg 87
 QY 61 tcttcacgcaaaagcgctctagccatgagcgtttagatagtgctgtgagcctccagcc 120
 DB 88 tcttcacgcaaaagcgctctagccatgagcgtttagatagtgctgtgagcctccagcc 147
 QY 121 cccccctccggagagacatagtgctgtcggaacgggtagtaacacggaattgcccag 180
 DB 148 cccccctccggagagacatagtgctgtcggaacgggtagtaacacggaattgcccag 207
 QY 181 gacgacgggtctcttcttgatcaatccgcctcaatgctcggaatttggcgctgccc 240
 DB 208 gacgacgggtctcttcttgatcaatccgcctcaatgctcggaatttggcgctgccc 266
 QY 241 cgcgagactgctagccgagtagtgctgtcggaacgggttagtaacacggaattgcccag 300
 DB 267 cgcgagactgctagccgagtagtgctgtcggaacgggttagtaacacggaattgcccag 326
 QY 301 ggtctgtgagagtgccccggagaggtctcgtagaccgtgacac 342
 DB 327 ggtctgtgagagtgccccggagaggtctcgtagaccgtgacac 368
 RESULT 6

E08872
 LOCUS E08872 4987 bp DNA
 DEFINITION DNA encoding a fusion protein of Vaccinia virus protein, HCV protein and Firefly luciferase.
 ACCSSION E08872

REFERENCE	JOURNAL	TITLE	AUTHORS	COMMENT
1 (bases 1 to 9431)				
1		Direct Submission		
		Submitted (14-JAN-1995) to the DDBJ/EMBL/Genbank databases. Makoto Seki, Mitsubishi Chemical Corporation Yokohama Research Center, Department of Bioscience Laboratory, 1000 Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227, Japan		
		(E-mail: seki.m@lilas.rc.m.kagaku.co.jp, Tel:045-963-3455, Fax:045-963-3992)		
2 (bases 1 to 9431)				
		2. (bases 1 to 9431)		
		Phosphorothioate antisense oligodeoxynucleotides capable of inhibiting hepatitis C virus gene expression: In vitro translation assay		
	JOURNAL	J. Biochem.	118 (6), 1199-1204 (1995)	
	1 MEDLINE	96362158		
	COMMENT	Sequence updated (03-May-1995) by: Makoto Seki.		
FEATURES				
source		Location/Qualifiers		
		1. 9431		
		/organism="Hepatitis C virus"		
		/db_xref="taxon:11103"		
		/clone="MKC1A"		
misc_feature		1. 2006		
		/note="clone N1-19 derived from the region was constructed clone N1-1, N3-1, N10-1, N27-3 and N19-1"		
		342. .9374		
		/codon_start=1		
		/product="HcV polypeptain"		
		/protein_id="BA08120.1"		
		/db_xref="GI:1212742"		
		/translation="MSTNPKPKRIKRNTRRPODYKFPGGQIVGGVLLIPRPRPLGVATRTKTSERPGRGROPPIKRAOGEQEGAPGMPFLYNEGIGAGMGLLSPGSRSPGQTPDRRRSRNLGKIDITLTCGALDLMKYIPLVAGPILGAAALAHGVRLDGVNATGONLPGCSFIFLLALISCLTIPASAYOYRNASGYHTINDCNSSITYEADVIMHTPCVPCVRENSSRCWVALPTLRANSSIPPTTIRRVLDLNGAALCSAYVGOEGSVPLVSOLFTEFSPRRYEVDOCSIYPGHVSGHRMADMIMNSPTTALVSQLRIPOAVLVDAVAGAHMVLAVGLAYSAVGMAKLVNMLFAGVDGGHTGKVAATGOTGTFEFSRGSOKIOLTNVNSMHIINRALNDSLITGLAALTYTSPNASCGERBAACGRPIDDEPOGMPILTHVNSINDORPCMHYAPRCGITYPAQVCGPYCCPTSPVYVGTTRDRAFTYTWGNEDVLLNNTPRPGMFGCTMNGTGFYKTCGGPCPNIGVGNNTLTCTPDCRKHPRHEAYTTCGSGPMVLTPLCLVHYPLMHPDYVNF TIKVRYVGGVEHRLBAACNMTRGECSDLEDNRSLSPLLSTEMOVLPCSFETLLPALSTGLIHLHQNIVDOYLYIGLSAVSAIKMEYILLFLPLADARVCAMLMMLLLIAOAEALENLYLVNAYSMAGAGILSFIEFPCAAMYIKRLVAGAYARVYWPULLLMALPRAYANDREMAASCGAAYFVCLVLTLSPPYKVLAKLIMQLITIRAEHLQYMIPLNRYGCRDAILITLCAVHPBLITDTIKLLAILGLPMVLOAGITRVYEVRAQGLIRCMKLVNRKAGGVYQMAFVKLAALITGYIYDHLPLRDMANALADLAAVEPVFSDMETKILTMGADVTACGDIIILGLPYVSARGKELILGPDLSLEGOMRLAPIT AYSOQTGGLGCIITSLTGRKNQVEEVOVSNANOSFLATCINGCWTVYAGASK TLAGKPIOTMTNTNDOLVGNPARGASMSPCDGGSDILVLRHADVITVYRRGDSGSLSPRITSLYKSSGGPLILCPSHGYIFGRVACVTRGAKANDVPYEMSET MRSVFNDSTPRAPVPOSFVAHLHABTGSKSTKPPAAIAGGYKVLNPSVAATLFGFAYMSKARGVDNPIRTGYRTTGAPIITYSYGKFLADGGSGCAVDITIDECSS TJDSTSLIGTIVDOAETAGARLVLTATAPGSAVYVPHNIEVALNSGELIPYK AILPEALIKGRHLIFCHSKCDELAALKLSALGVNAYAYRGDVSIIPTSGVWVVA TDLMLGYTGDPSVDICNCTCVOTDESLDPTFTTETTPODVARSORSGRTVGR RGLIETRVTPGERSGCMFDSYVCECDACGACAYELTPATSYRLAYLNTPELPCO DDLFEMSVTGLTIHDHAFLSOTKQKQNDPVLVAQIATCARAKAPSPWOMMCL LIRKPLHGPRTPLYLRLGAVONEVLTPLITF IMACMSADLEVIVSTVLAQVGLA ALAAKCLTTSQSVIVGRILISGRPAVLPDEVLVYOEDEMEBCASHLPYIEQOMLPE OFOKALGILGLOTATROKAEAAAPVYESKRALTEFKAKHMMNTISGLOYLGLTPLEN PAIISAKMPTFASITPTTOYTLITNLGFWGMAVOLAPRASAAPNAGAGIAGAVYCI GLCKVLYDLIASIGAYAGALVAFKVMSCMPSTEDLVNLPILISGALVYVCAAILRRHVGEGGAOVOMNRLIAFASRGHNSPTHYVPSDDAARVTOILSLITITQLKRLHQMIDEDSTPCSGSLMDVDWIMICTVLADTKTOSMLILRPLCPVPFSCORGYK GWRGQDIYNTTICPGCAOLIGHWKNGSMIRLPTSCSNTHMGFPIIAUYTIGCTPSP APNYSRLMVAABEYEVTRGDFHVTGEMTDNPKCQVYVAPEFTLIDCARLRLR YAPACKRLADEVTFQVGNQYVSGSLPCEBPBDVYVSMPLDEPSHTIADARRL ARSPPRLASSSSQLSALSKATCTTHHGAPTDLEANLIRORMGVITRYESN KIVILDSFELRLAEEDEEVVAAILRTKRPKPAAPVARDYVNPILLSKANDPYVPVILHSCPLPPKRAPIPPPRKRRKIVTLVLETSVSSALALARTKTSGSSSAVDST		

Query Match 95.6%; Score 326.8; DB 14; Length 9431;
 Best Local Similarity 99.1%; Pred. No. 3.5e-82;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 gccagcccccctgtaggggagacacccacacataagatcccccgtgtaggagactacg 60
 DB 1 gccagcccccctgtaggggagacacccacacataagatcccccgtgtaggagactacg 60
 QY 61 tcttcacgagaaagcgtctagccatgagcgttagtagtgcgtgcagacccccaagcc 120
 DB 61 tcttcacgagaaagcgtctagccatgagcgttagtagtgcgtgcagacccccaagcc 120
 QY 121 ccccccctcccgagagagacatagtgctgcggaacccgtgtagtacacggaaatgcag 180
 DB 121 ccccccctcccgagagagacatagtgctgcggaacccgtgtagtacacggaaatgcag 180
 QY 181 gacgacccgggtctctcttgatcacaatcccgctcaatgctgtagaattgggggtgcccc 240
 DB 181 gacgacccgggtctctcttgatcacaatcccgctcaatgctgtagaattgggggtgcccc 240
 QY 241 cgcgagacgtctagccgagtagtgcgtgcggaagacccctgtagtactgcctagtag 300
 DB 241 cgcgagacgtctagccgagtagtgcgtgcggaagacccctgtagtactgcctagtag 300
 QY 301 ggtgcttgcgagtgcccgaggaggtctcgttagacgctgcac 342
 DB 301 ggtgcttgcgagtgcccgaggaggtctcgttagacgctgcac 342
 Db 300 ggtgcttgcgagtgcccgaggaggtctcgttagacgctgcac 341

RESULT 9

E06261 9471 bp RNA PAT 29-SEP-1997
 LOCUS E06261 CDNA encoding genes derived from hepatitis C virus.
 DEFINITION E06261.1 GI:2174448
 ACCESSION E06261.1 GI:2174448
 VERSION JP 1994000085-A/101.
 KEYWORDS Hepatitis C virus.
 SOURCE Hepatitis C virus.
 ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1 (bases 1 to 9471)
 AUTHORS Seki,M., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N..
 GENE OR DNA FRAGMENT DERIVED FROM HEPATITIS C VIRUS, POLYPEPTIDE CODED BY THE SAME AND ITS PRODUCTION
 JOURNAL Patent: JP 1994000085-A 101 11-JAN-1994;
 MITSUBISHI KASEI CORP
 COMMENT OS (hepatitis C virus)
 PN JP 1994000085-A/101

PD 11-JAN-1994
 PR 11-JUN-1992 JP 1992194497
 PR 11-JUN-1991 JP 91P 139268, 12-JUL-1991 JP 91P 172794, PR
 07-OCT-1991 JP 91P 287008, 16-DEC-1991 JP 91P 332329, PR
 20-APR-1992 JP 92P 99957
 PI SEKI MAKOTO, HONDA YOSHIKAZU, TAKAHASHI KAZUNOBU, PI
 MURAKAMI TOMOKO,
 PI TERANISHI YUTAKA, HAYASHI NORIO
 PC C12N15/51,C07K7/06,C07K7/08,C07K13/00,C07K15/12, PC
 C12N1/21,C12N5/10,
 PC C12N15/11,C12N15/70,C12N15/85,C12P21/02//A61K9/00,A61K9/29,
 PC C12N1/21,
 PC C12R1/19,(C12N5/10,C12R1/91),(C12P21/02,C12R1/19),(C12P21/02,
 PC C12R1/91),
 PC C07K9/00;
 CC strandedness: Double;
 CC topology: Linear;
 CC anti-sense: No;
 CC *source: clone-T7N1-30;
 FH Key Location/Qualifiers
 FT mat_peptide 369..9398
 FT /product- the peptides reacting specifically

FT and
 FT immunochemically with the serum of hepatitis
 FT type C
 FT patient'.
 FT

FEATURES

source 1..9471 Location/Qualifiers
 /organism="Hepatitis C virus"
 /db_xref="taxon:11103"

BASE COUNT 1892 a 2830 c 2712 g 2037 t
 ORIGIN

Query Match 95.6%; Score 326.8; DB 6; Length 9471;
 Best Local Similarity 99.1%; Pred. No. 3.5e-82;
 Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 gccagcccccctgtaggggagacacccacacataagatcccccgtgtaggagactacg 60
 DB 28 gccagcccccctgtaggggagacacccacacataagatcccccgtgtaggagactacg 87
 QY 61 tcttcacgagaaagcgtctagccatgagcgttagtagtgcgtgcagacccccaagcc 120
 DB 88 tcttcacgagaaagcgtctagccatgagcgttagtagtgcgtgcagacccccaagcc 147
 QY 121 ccccccctcccgagagagacatagtgctgcggaacccgtgtagtacacggaaatgcag 180
 DB 148 ccccccctcccgagagagacatagtgctgcggaacccgtgtagtacacggaaatgcag 207
 QY 181 gacgacccgggtctctcttgatcacaatcccgctcaatgctgtagaattgggggtgcccc 240
 DB 208 gacgacccgggtctctcttgatcacaatcccgctcaatgctgtagaattgggggtgcccc 266
 QY 241 cgcgagacgtctagccgagtagtgcgtgcggaagacccctgtagtactgcctagtag 300
 DB 257 cgcgagacgtctagccgagtagtgcgtgcggaagacccctgtagtactgcctagtag 326
 QY 301 ggtgcttgcgagtgcccgaggaggtctcgttagacgctgcac 342
 DB 327 ggtgcttgcgagtgcccgaggaggtctcgttagacgctgcac 368

RESULT 10

E06457 9471 bp RNA PAT 29-SEP-1997
 LOCUS E06457 CDNA encoding genes derived from hepatitis C virus.
 DEFINITION E06457
 ACCESSION E06457.1 GI:2174644
 VERSION JP 1994000086-A/101.
 KEYWORDS Hepatitis C virus.
 SOURCE Hepatitis C virus.
 ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1 (bases 1 to 9471)
 AUTHORS Seki,M., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N..
 GENE FOR DNA FRAGMENT DERIVED FROM HEPATITIS C VIRUS, POLYPEPTIDE CODED BY THE SAME AND ITS PRODUCTION
 JOURNAL Patent: JP 1994000086-A 101 11-JAN-1994;
 MITSUBISHI KASEI CORP
 COMMENT OS (hepatitis C virus)
 PN JP 1994000086-A/101
 PD 11-JAN-1994
 PR 07-OCT-1992 JP 1992293734
 PR 07-OCT-1991 JP 91P 287008, 16-DEC-1991 JP 91P 332329, PR
 20-APR-1992 JP 92P 99957
 PI SEKI MAKOTO, HONDA YOSHIKAZU, TAKAHASHI KAZUNOBU, PI
 MURAKAMI TOMOKO,
 PI TERANISHI YUTAKA, HAYASHI NORIO
 PC C12N15/51,C07K7/06,C07K7/08,C07K13/00,C12N5/10, PC
 C12N15/11,
 PC C12N15/85,C12P21/02//A61K9/29,(C12P21/02,C12R1/91),C07K9/00;
 CC strandedness: Double;
 CC topology: Linear;

CC anti-sense: No;
CC *source: clone=T7N1-30;
FH key Location/Qualifiers
FT mat_peptide 369..9398
FT /product='the peptides reacting specifically
FT and
FT type C
FT immunologically with the serum of hepatitis
FT patient'.
FEATURES
source Location/Qualifiers
1..9471
/organism="Hepatitis C virus"
/db_xref="taxon:11103"
BASE COUNT 1892 a 2830 c 2712 g 2037 t
ORIGIN

Query Match 95.1%; Score 326.8; DB 6; Length 9471;
Best Local Similarity 99.1%; Pred. No. 3.5e-82;
Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 gccagccccctgattgaggcgacactccacatagatcctccctgtgaggaactactg 60
DB gccagccccctgattgaggcgacactccacatagatcctccctgtgaggaactactg 87
QY 61 tcttacgcagaagaagcgtctagatcctgagtagatgagtgctgtgcaagcctcagacc 120
DB tcttacgcagaagaagcgtctagatcctgagtagatgagtgctgtgcaagcctcagacc 147
QY 88 tcttacgcagaagaagcgtctagatcctgagtagatgagtgctgtgcaagcctcagacc 180
DB tcttacgcagaagaagcgtctagatcctgagtagatgagtgctgtgcaagcctcagacc 207
QY 121 ccccccctccgagagacatagtgctgaggaacccgagtagacacccgaggaattgcag 180
DB ccccccctccgagagacatagtgctgaggaacccgagtagacacccgaggaattgcag 207
QY 148 ccccccctccgagagacatagtgctgaggaacccgagtagacacccgaggaattgcag 240
DB ccccccctccgagagacatagtgctgaggaacccgagtagacacccgaggaattgcag 266
QY 208 gacgacccgggtctcttctgagtaaa-cccgctcaatgcctgaggaattgagcgccccc 240
DB gacgacccgggtctcttctgagtaaa-cccgctcaatgcctgaggaattgagcgccccc 266
QY 241 cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 300
DB cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 326
QY 267 cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 342
DB cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 368
QY 301 ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342
DB ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 368
QY 327 ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 368
DB ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 368

RESULT 11
ARI19856 341 bp DNA PAT 16-MAY-2001
LOCUS Sequence 48 from patent US 6153421.
DEFINITION ARI19856
ACCESSION ARI19856
VERSION ARI19856.1 GI:14102555
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 341)
AUTHORS Yanagi,M., Bukh,J., Emerson,S.U. and Purcell,R.H.
TITLE Cloned genomes of infectious hepatitis C viruses and uses thereof
JOURNAL Patent: US 6153421-A 48 28-NOV-2000;
FEATURES
source Location/Qualifiers
1..341
/organism="unknown"
BASE COUNT 63 a 103 c 106 g 69 t
ORIGIN

Query Match 95.1%; Score 325.2; DB 6; Length 341;
Best Local Similarity 98.8%; Pred. No. 1.7e-81;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagccccctgattgaggcgacactccacatagatcctccctgtgaggaactactg 60
DB gccagccccctgattgaggcgacactccacatagatcctccctgtgaggaactactg 60

DB gccagccccctgattgaggcgacactccacatagatcctccctgtgaggaactactg 60
QY 61 tcttacgcagaagaagcgtctagatcctgagtagatgagtgctgtgcaagcctcagacc 120
DB tcttacgcagaagaagcgtctagatcctgagtagatgagtgctgtgcaagcctcagacc 120
QY 121 ccccccctccgagagacatagtgctgaggaacccgagtagacacccgaggaattgcag 180
DB ccccccctccgagagacatagtgctgaggaacccgagtagacacccgaggaattgcag 180
QY 181 gacgacccgggtctcttctgagtaaa-cccgctcaatgcctgaggaattgagcgccccc 240
DB gacgacccgggtctcttctgagtaaa-cccgctcaatgcctgaggaattgagcgccccc 240
QY 241 cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 300
DB cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 300
QY 267 cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 342
DB cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 342
QY 301 ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342
DB ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342
QY 327 ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342
DB ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342

RESULT 12
ARO30378 9401 bp DNA PAT 29-SEP-1999
LOCUS Sequence 1 from patent US 5861267.
DEFINITION ARO30378
ACCESSION ARO30378
VERSION ARO30378.1 GI:5943592
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9401)
AUTHORS Su,M.
TITLE Methods, nucleotide sequences and host cells for assaying exogenous
and endogenous protease activity
JOURNAL Patent: US 5861267-A 1 19-JAN-1999;
FEATURES
source Location/Qualifiers
1..9401
/organism="unknown"
BASE COUNT 1868 a 2848 c 2686 g 1999 t
ORIGIN

Query Match 95.1%; Score 325.2; DB 6; Length 9401;
Best Local Similarity 98.8%; Pred. No. 1e-81;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagccccctgattgaggcgacactccacatagatcctccctgtgaggaactactg 60
DB gccagccccctgattgaggcgacactccacatagatcctccctgtgaggaactactg 60
QY 61 tcttacgcagaagaagcgtctagatcctgagtagatgagtgctgtgcaagcctcagacc 120
DB tcttacgcagaagaagcgtctagatcctgagtagatgagtgctgtgcaagcctcagacc 120
QY 121 ccccccctccgagagacatagtgctgaggaacccgagtagacacccgaggaattgcag 180
DB ccccccctccgagagacatagtgctgaggaacccgagtagacacccgaggaattgcag 180
QY 181 gacgacccgggtctcttctgagtaaa-cccgctcaatgcctgaggaattgagcgccccc 240
DB gacgacccgggtctcttctgagtaaa-cccgctcaatgcctgaggaattgagcgccccc 240
QY 241 cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 300
DB cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 300
QY 267 cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 342
DB cgcgagactgctgagcagagtagtggtgaggaagcctgtgtagcctcagtag 342
QY 301 ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342
DB ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342
QY 327 ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342
DB ggtgcttgagagtgcccgagaggtctcgtagacagctgacac 342

Db 300 GGAGCTTGGAGTCCCGGAGGAGCTGCTGAGACCGGTGACC 341

QY	1	gcacgccccctgtagggggcgacacctccacatgatacaaccccccttggagaaactctg	60
Db	1	GCACGCCCCCTGTAGGGGGCGACACTCCACATGATCAACCCCCCTTGGAGAACTCTG	60
QY	61	tcttcagcgagaaagcgctctagccatgycgtctagtatgagtgtctgcagcctccagcc	120
Db	61	TCTTTCAGCGAAGAAAGCGCTCTAGCCATG6CGTTAGATATGAGTGTCTGCACCTCCAGGAC	120
QY	121	ccccctctccgggagagccatagtgctgcgcgaaacggctgagtaacacggaaattgcag	180
Db	121	CCCCCTCTCCGGGAGAGCCATAGTGGTCTGGGGAACCGGTGATGACACCGGAATTGCCAG	180
QY	181	gagacacgggagccttctcttgatacaatcccgctcaatgacctgagattggcggtgcccc	240
Db	181	GACGACCGGGTCTCTTCTTGAGATCA-CCGCGCTAAAGCCCTGGAGATTGGCGTGCACC	229
QY	241	cgcgagactgctagccgagtagtgcttggtcgcgaaagcctctggtctacgctgtag	300
Db	240	CGCGAGACTGCTTACCCGACAGTGTGTGGGTCCGCAAAAGCGCTTGCTACTGCCTGATAG	299
QY	301	ggtgcttgagagtgcccccggaagcttcgcgtagaaacgcygaccc	342
Db	300	GGTGCTTGCAGATGCCCGCGGAGGTCTCGTAGACCGTGCACATC	341

RESULT 15
AR119832

LOCUS	AR119832	9595 bp	DNA	PAT	16-MAY-2001
DEFINITION	Sequence 4 from patent US 6153421.				

ACCESSION	AR119832
VERSION	AR119832.1
REVISION	GI:14102531

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE	BASE COUNT	ORIGIN
Unknown.	Unknown.	Unknown.	Unknown.	Unclassified.	1 (bases 1 to 9595)	1934	a 2842 c 2698 g 2121 t
Yanai,M., Bukh,J., Emerson,S.U. and Purcell,R.H.	Cloned genomes of infectious hepatitis C viruses and uses thereof	Patent: US 6153421-A 4 28-NOV-2000;	location/Qualifiers	1..9595	/organism="unknown"	1934	a 2842 c 2698 g 2121 t

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

PS Claim 24; Page 64; 94pp; Japanese.

XX The present invention describes an oligonucleotide sequence which
 CC potentiates the expression of a useful gene when incorporated in a
 CC gene expression vector, where the oligonucleotide sequence consists
 CC of the 5'-non-translated region (5'UTR) of a viral genome or its
 CC fragment or modified form. The 5'UTR sequence is useful in screening
 CC potential initiation factors of interaction with the viral IRES or
 CC IRES-dependency, for treatment of diseases caused by cap-dependent
 CC mRNA translation, and in the diagnosis of the severity of HCV infection.
 CC The present sequence represents a mutated Hepatitis type C virus (HCV)
 CC 5'UTR nucleotide sequence, which is used in the exemplification of
 CC the present invention.

XX Sequence 342 BP; 62 A; 104 C; 106 G; 70 T; 0 other;

Query Match 100.0%; Score 342; DB 21; Length 342;
 Best Local Similarity 100.0%; Pred. No. 3.3e-92;
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gccagcccccctgattgaggcgacactccacacatagatcaccctctgaggaactactg 60
 DB 1 gccagcccccctgattgaggcgacactccacacatagatcaccctctgaggaactactg 60
 OY 61 tcttacgagaagaacgctctagcattgagtgatgagtgctgagcagctccagagcc 120
 DB 61 tcttacgagaagaacgctctagcattgagtgatgagtgctgagcagctccagagcc 120
 OY 121 ccccccctccgggagagacatagtgctcggaacccggtgagacacgggaattggcag 180
 DB 121 ccccccctccgggagagacatagtgctcggaacccggtgagacacgggaattggcag 180
 OY 181 gacgacccggtctctctcttgatcaatcccgctcaatgcttgagattggcgctgccc 240
 DB 181 gacgacccggtctctctcttgatcaatcccgctcaatgcttgagattggcgctgccc 240
 OY 241 cgcgagactgctagccgagtagtgctgggtcggaagagccttggtactgctcgatag 300
 DB 241 cgcgagactgctagccgagtagtgctgggtcggaagagccttggtactgctcgatag 300
 OY 301 ggtgctgcgagtgcccgaggagctcgtagaccgtgcacc 342
 DB 301 ggtgctgcgagtgcccgaggagctcgtagaccgtgcacc 342

RESULT 2

AAA08097
 ID AAA08097 standard; cDNA: 713 BP.

XX AAA08097;

XX 22-JUN-2000 (first entry)

XX Hepatitis type C virus nucleotide sequence SEQ ID NO:1.

XX Hepatitis C virus; HCV; 5' UTR; 5'-nontranslational region; diagnosis;
 KW gene expression; infection; IRES; viral; ss.

OS Hepatitis C virus.

XX Key Location/Qualifiers

FT 5'UTR 1..341

FT CDS /*tag= a
 /*tag= b
 /note= "no stop codon given"

XX WO200012691-A1.

XX 09-MAR-2000.

XX 08-JUL-1999; 99WO-JP03682.

XX 27-AUG-1998; 98JP-0241367.

XX (EJSO) FUSO PHARM IND LTD.

XX Yamada O, Yoshida H, Zhang J;

XX WEI; 2000-237867/20.

XX P-PSDB; AAV82358.

XX Nucleic acid sequence which potentiates gene expression and contains
 PT 5'-non-translated region of a viral genome is incorporated in vectors
 PT for improving the expression of a useful gene

XX Claim 18; Page 82-84; 94pp; Japanese.

XX The present invention describes an oligonucleotide sequence which
 CC potentiates the expression of a useful gene when incorporated in a
 CC gene expression vector, where the oligonucleotide sequence consists
 CC of the 5'-non-translated region (5'UTR) of a viral genome or its
 CC fragment or modified form. The 5'UTR sequence is useful in screening
 CC potential initiation factors of interaction with the viral IRES or
 CC IRES-dependency, for treatment of diseases caused by cap-dependent
 CC mRNA translation, and in the diagnosis of the severity of HCV infection.
 CC The present sequence represents a Hepatitis type C virus (HCV) nucleotide
 CC sequence comprising a 5'UTR, which is used in the exemplification of
 CC the present invention.

XX Sequence 713 BP; 136 A; 218 C; 226 G; 133 T; 0 other;

Query Match 96.0%; Score 328.4; DB 21; Length 713;
 Best Local Similarity 99.4%; Pred. No. 4.3e-88;
 Matches 340; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 gccagcccccctgattgaggcgacactccacacatagatcaccctctgaggaactactg 60
 DB 1 gccagcccccctgattgaggcgacactccacacatagatcaccctctgaggaactactg 60
 OY 61 tcttacgagaagaacgctctagcattgagtgatgagtgctgagcagctccagagcc 120
 DB 61 tcttacgagaagaacgctctagcattgagtgatgagtgctgagcagctccagagcc 120
 OY 121 ccccccctccgggagagacatagtgctcggaacccggtgagacacgggaattggcag 180
 DB 121 ccccccctccgggagagacatagtgctcggaacccggtgagacacgggaattggcag 180
 OY 181 gacgacccggtctctctcttgatcaatcccgctcaatgcttgagattggcgctgccc 240
 DB 181 gacgacccggtctctctcttgatcaatcccgctcaatgcttgagattggcgctgccc 240
 OY 241 cgcgagactgctagccgagtagtgctgggtcggaagagccttggtactgctcgatag 300
 DB 241 cgcgagactgctagccgagtagtgctgggtcggaagagccttggtactgctcgatag 300
 OY 301 ggtgctgcgagtgcccgaggagctcgtagaccgtgcacc 342
 DB 301 ggtgctgcgagtgcccgaggagctcgtagaccgtgcacc 341

RESULT 3

AA064913
 ID AA064913 standard; DNA; 2033 BP.

XX AA064913;

XX 21-DEC-1994 (first entry)

XX Hepatitis C virus genome partial sequence (5'-end).

XX Hepatitis C virus; Non-A, non-B hepatitis virus; HCV; antisense;
 KW therapy; inhibition; viral protein precursor; core protein;
 KW envelope protein; non-structural protein; clone 17NI-19; ss.

```
XX Hepatitis C virus.
OS ID A086788 standard; DNA; 2033 BP.
XX AC A086788;
XX FH Key Location/Qualifiers
XX FT CDS 369..2033
XX FT /tag=a
XX FT /note="partial open reading frame"
XX
XX CA2104649-A.
XX
XX 26-FEB-1994.
XX
XX 23-AUG-1993; 93CA-2104649.
XX
XX 25-AUG-1992; 92JP-0248796.
XX PR 03-MAR-1993; 93JP-0042736.
XX
XX (SEKI/) SEKI M.
XX
XX Honda Y, Seki M, Yamada E;
XX
XX WPI; 1994-151836/19.
XX DR P-PSDB; AAR54866.
XX
XX Anti-sense oligo:nucleotide(s) complementary to the hepatitis C
XX PT virus genome - are useful as antiviral agents
XX
XX Claim 1; Page 59-64; 262pp; English.
XX
XX Claimed antisense oligonucleotides have a sequence complementary to
XX a base sequence which consists of 10-34 bases, extracted from:
XX (1) 93 bases from T at position 107 to A at position 199 of A064913,
XX (2) 152 bases from A at position 250 to C at position 401 of
XX A064913, or (3) 52 bases from C at position 808 to A at position 859
XX of A064913. The HCV sequence A064913 comprises the 5'-end of the
XX viral open reading frame, including the core protein coding
XX sequence. The new antisense oligonucleotides are useful for
XX inhibiting translation of the single polypeptide precursor which is
XX subsequently cleaved to produce the HCV viral proteins.
XX
XX Sequence 2033 BP; 379 A; 612 C; 595 G; 447 T; 0 other;
XX
XX Query Match 95.6%; Score 326.8; DB 15; Length 2033;
XX Best Local Similarity 99.1%; Pred. No. 1.7e-87;
XX Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
XX
XX 1 gccagccctctgtagggcgacacccacacatagacacccctctgtaggaactactg 60
XX 28 gccagccctctgtagggcgacacccacacatagacacccctctgtaggaactactg 87
XX
XX 61 tcttcacgagaagaagcgtctagcattagtagtgcgtgcgcagccctccagacc 120
XX 88 tcttcacgagaagaagcgtctagcattagtagtgcgtgcgcagccctccagacc 147
XX
XX 121 ccccccctccggagagacatagtgctgcgcgaacccgtgagttacacccgaattgcag 180
XX 148 ccccccctccggagagacatagtgctgcgcgaacccgtgagttacacccgaattgcag 207
XX
XX 181 gccagcccggtctctctctggaatcccgctcaatgacctgagattgggggtgcccc 240
XX 208 gccagcccggtctctctctggaatcccgctcaatgacctgagattgggggtgcccc 266
XX
XX 241 ccgagacactgctagccgagtagtgcgtgcgcgaacccgtgtagtgcctatag 300
XX 267 ccgagacactgctagccgagtagtgcgtgcgcgaacccgtgtagtgcctatag 326
XX
XX 301 ggtgctgcgagtgcccgaggaggtctcgtagacgcgtgcaac 342
XX 327 ggtgctgcgagtgcccgaggaggtctcgtagacgcgtgcaac 368
XX
XX RESULT 4
```

```
A086788
XX ID A086788 standard; DNA; 2033 BP.
XX AC A086788;
XX FH Key Location/Qualifiers
XX FT CDS 369..2033
XX FT /tag=a
XX FT /note="STOP codon absent"
XX
XX JP07069899-A.
XX
XX 14-MAR-1995.
XX
XX 02-SEP-1993; 93JP-0241973.
XX PR 02-SEP-1993; 93JP-0241973.
XX
XX (MITSUBISHI KASEI CORP.
XX
XX WPI; 1995-144713/19.
XX DR P-PSDB; AAR72800.
XX
XX Antiviral agent comprising component which disrupts viral gene
XX PT translation - used for the selective inhibition of e.g. Hepatitis
XX C virus, polio-virus and human rhinovirus
XX
XX Disclosure; Pages 12-14; 23pp; Japanese.
XX
XX A086788 encodes AAR72800 Hepatitis C virus (HCV) T7NI-19, which
XX CC disrupts viral gene translation, by preventing the binding of
XX CC the viral mRNA to the internal ribosome entry site. It was
XX CC used in the construction of an antiviral agent detecting vector,
XX CC which comprises a vaccinia virus vector, a vector containing
XX CC T7NI-19 and a firefly luciferase gene. The antiviral agent can
XX CC be used to treat HCV, poliovirus, cerebral cardio-hepatitis,
XX CC human rhinovirus and foot and mouth disease viral infections.
XX
XX Sequence 2033 BP; 379 A; 612 C; 595 G; 447 T; 0 other;
XX
XX Query Match 95.6%; Score 326.8; DB 16; Length 2033;
XX Best Local Similarity 99.1%; Pred. No. 1.7e-87;
XX Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
XX
XX 1 gccagccctctgtagggcgacacccacacatagacacccctctgtaggaactactg 60
XX 28 gccagccctctgtagggcgacacccacacatagacacccctctgtaggaactactg 87
XX
XX 61 tcttcacgagaagaagcgtctagcattagtagtgcgtgcgcagccctccagacc 120
XX 88 tcttcacgagaagaagcgtctagcattagtagtgcgtgcgcagccctccagacc 147
XX
XX 121 ccccccctccggagagacatagtgctgcgcgaacccgtgagttacacccgaattgcag 180
XX 148 ccccccctccggagagacatagtgctgcgcgaacccgtgagttacacccgaattgcag 207
XX
XX 181 gccagcccggtctctctctggaatcccgctcaatgacctgagattgggggtgcccc 240
XX 208 gccagcccggtctctctctggaatcccgctcaatgacctgagattgggggtgcccc 266
XX
XX 241 ccgagacactgctagccgagtagtgcgtgcgcgaacccgtgtagtgcctatag 300
XX 267 ccgagacactgctagccgagtagtgcgtgcgcgaacccgtgtagtgcctatag 326
XX
XX 301 ggtgctgcgagtgcccgaggaggtctcgtagacgcgtgcaac 342
XX 327 ggtgctgcgagtgcccgaggaggtctcgtagacgcgtgcaac 368
XX
XX RESULT 4
```

Db 267 cgcgagactgctagccgagtagtggtggtcgagaaagccttggtgctgctgcatagatg 326
QY 301 ggtgcttcggaagtcgcccgaggagctctcgttagaccgtgcacc 342
|||||
Db 327 ggtgcttcggaagtcgcccgaggagctcgttagaccgtgcacc 368
|||||
RESULT 5
AA065322
ID AA065322 standard; DNA; 4987 BP.
XX
AC AA065322;
XX
DT 21-DEC-1994 (first entry)
XX
DE Vaccinia virus promoter and HCV core protein-luciferase fusion gene.
XX
KW Hepatitis C Virus; HCV; control; Non-A, non-B hepatitis virus;
KW antisense; therapy; inhibition; viral protein precursor;
KW recombinant vaccinia virus; HCV core protein gene; firefly;
KW luciferase reporter gene; fusion construct; ds.
XX
OS Chimeric Vaccinia virus.
OS Chimeric Hepatitis C virus.
OS Chimeric Photinus pyralis.
XX
FH Key Location/Qualifiers
FT CDS 1826..4057
FT /*tag= a
FT /note= "encodes HCV core protein-luciferase
fusion protein"
XX
PN CA2104649-A.
XX
PD 26-FEB-1994.
XX
PF 23-AUG-1993; 93CA-2104649.
XX
PR 25-AUG-1992; 92JP-0248796.
PR 03-MAR-1993; 93JP-0042736.
XX
PA (SEKI/) SEKI M.
XX
PI Honda Y, Seki M, Yamada E;
XX
DR WPI: 1994-151836/19.
DR P-PSDB: AAR34867.
XX
PT Anti-sense oligo:nucleotide(s) complementary to the hepatitis C
PT virus genome - are useful as antiviral agents
XX
PS Example 5; Page 243-251; 262pp; English.
XX
CC A recombinant vaccinia virus which codes for a HCV core protein
CC fused to the firefly luciferase enzyme was constructed from PCR
CC amplified fragments. The construct was useful for assaying the
CC inhibitory activity of various antisense oligonucleotides on HCV
CC gene translation.
XX
SQ Sequence 4987 BP; 1520 A; 1052 C; 1083 G; 1332 T; 0 other;
XX
QY Query Match 95.6%; Score 326.8; DB 15; Length 4987;
Best Local Similarity 99.1%; Pred. No. 2.1e-87;
Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Db 1485 gcccgcgcctgagtgaggagcaccacacacatagatcaccctctgaggaactactg 60
|||||
QY 61 tcttcacgcagaaagcgtcttagcattgagtggtggtgagcctccagcc 120
|||||
Db 1545 tcttcacgcagaaagcgtcttagcattgagtggtggtgagcctccagcc 1604
|||||

QY 121 ccccccctccgggagagacatagtgctcgcgaaaccggtgagtaaccggaaattgcag 180
|||||
Db 1605 ccccccctccgggagagacatagtgctcgcgaaaccggtgagtaaccggaaattgcag 1664
|||||
QY 181 gacgacgcggtctcttcttgatcaatcccgctcaatgctggaattggcgctgcc 240
|||||
Db 1665 gacgacgcggtctcttcttgatcaatcccgctcaatgctggaattggcgctgcc 1723
|||||
QY 241 cgcgagactgctagccgagtagtggtggtcgagaaagccttggtgctgctgcatag 300
|||||
Db 1724 cgcgagactgctagccgagtagtggtggtcgagaaagccttggtgctgctgcatag 1783
|||||
QY 301 ggtgcttcggaagtcgcccgaggagctcgttagaccgtgcacc 342
|||||
Db 1794 ggtgcttcggaagtcgcccgaggagctcgttagaccgtgcacc 1825
|||||
RESULT 6
AA086799
ID AA086799 standard; DNA; 4987 BP.
XX
AC AA086799;
XX
DT 04-DEC-1995 (first entry)
XX
DE Vaccinia virus vector comprising HCV T7N1-19 and firefly luciferase.
XX
KW Vaccinia virus vector; firefly luciferase; hepatitis C virus; T7N1-19;
KW antiviral agent; poliovirus; human rhinovirus;
KW internal ribozyme entry site; non-A non-B; cerebral cardio-hepatitis;
KW foot and mouth disease; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1826..4057
FT /*tag= a
FT /*tag= a
XX
PN JP07069899-A.
XX
PD 14-MAR-1995.
XX
PF 02-SEP-1993; 93JP-0241973.
XX
PR 02-SEP-1993; 93JP-0241973.
XX
PA (MITU) MITSUBISHI KASEI CORP.
XX
DR WPI: 1995-144713/19.
DR P-PSDB: AAR72801.
XX
PT Antiviral agent comprising component which disrupts viral gene
PT translation - used for the selective inhibition of e.g. Hepatitis
PT C virus, polio:virus and human rhinovirus
XX
PS Disclosure; Pages 19-22; 23pp; Japanese.
XX
CC AA086788 encodes AAR72800 Hepatitis C virus (HCV) T7N1-19, which
CC disrupts viral gene translation, by preventing the binding of
CC the viral mRNA to the internal ribozyme entry site. It was
CC used in the construction of an antiviral agent detecting vector
CC (AA086799 which encodes AAR72801), which comprises a vaccinia virus
CC vector, a vector containing T7N1-19 and a firefly luciferase gene.
CC The antiviral agent can be used to treat HCV, poliovirus, cerebral
CC cardio-hepatitis, human rhinovirus and foot and mouth disease viral
CC infections.
XX
SQ Sequence 4987 BP; 1516 A; 1050 C; 1081 G; 1322 T; 18 other;
XX
QY Query Match 95.6%; Score 326.8; DB 16; Length 4987;
Best Local Similarity 99.1%; Pred. No. 2.1e-87;

PN WO9634976-A1.
 XX 07-NOV-1996.
 PD
 XX 01-MAY-1996; 96WO-US06070.
 PF
 XX 01-MAY-1995; 95US-0432693.
 PR
 XX (VERT-) VERTEX PHARM INC.
 PA
 XX Su M;
 PI
 XX WPI: 1996-506176/50.
 DR
 XX Assay for protease activity, esp. viral protease - using host cells
 PT which secrete exogenous or endogenous proteases and an artificial
 PP polypeptide substrates
 PS
 XX Example 1; Page 21-25; 44pp; English.
 CC A full-length hepatitis C virus (HCV) H strain cDNA (AAT41882)
 CC includes regions coding for the NS3 and NS4 proteases. NS3 is a
 CC 70 kDa protease involved in the maturation of viral polypeptides
 CC following infection. A novel method for assaying protease (e.g.
 CC NS3) activity involves transforming a host cell with DNA encoding
 CC the protease and DNA encoding a substrate (e.g. interleukin-1 beta).
 CC When these nucleotide sequences are expressed, the protease cleaves
 CC the substrate, which is secreted from the host where it can be
 CC easily quantitated. The method can be used to identify potential
 CC inhibitors of the protease for therapeutic appln.
 CC
 XX
 SQ Sequence 9401 BP; 1868 A; 2848 C; 2686 G; 1999 T; 0 other;

Query Match 95.1%; Score 325.2; DB 17; Length 9401;
 Best Local Similarity 98.8%; Pred. No. 7.3e-87;
 Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagccccctgattgaggcgagacacccacatagatcctccctgtgaggaactactg 60
 DB 1 gccagccccctgattgaggcgagacacccacatagatcctccctgtgaggaactactg 60
 QY 61 tcttaacgagaagaagcgcttaacatgagcgtttagatgagtgctgacgctccagagcc 120
 DB 61 tcttaacgagaagaagcgcttaacatgagcgtttagatgagtgctgacgctccagagcc 120
 QY 121 ccccccctccgggagagacatagtgctgaggaacccgtgtagtaacccggaattgcag 180
 DB 121 ccccccctccgggagagacatagtgctgaggaacccgtgtagtaacccggaattgcag 180
 QY 181 gagcagcggggtctcttcttgatcaatccgcctcaatgctgagattggcggtccccc 240
 DB 181 gagcagcggggtctcttcttgatcaatccgcctcaatgctgagattggcggtccccc 240
 QY 241 cgcgagactgctcagcagatgctgtggttcgcaaaagcctctgtgactccgatatg 300
 DB 241 cgcgagactgctcagcagatgctgtggttcgcaaaagcctctgtgactccgatatg 300
 QY 301 ggtgctgcgagtgcccgcgagaggtctcgttagaccgtgcacc 342
 DB 301 ggtgctgcgagtgcccgcgagaggtctcgttagaccgtgcacc 342
 QY 342 ggtgctgcgagtgcccgcgagaggtctcgttagaccgtgcacc 341
 DB 342 ggtgctgcgagtgcccgcgagaggtctcgttagaccgtgcacc 341

RESULT 9
 ID AAX24843
 AAX24843 standard; DNA; 9595 BP;
 AC AAX24843;
 XX 21-JUN-1999 (first entry)
 DT
 XX Infectious hepatitis C virus genotype 1b strain HC-J4 genome.
 DE
 XX

KW HCV; Infectious clone; infection; diagnosis; therapy; vaccine;
 KW screening; assay; antiviral; virucide; ss.
 XX
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 342..9374
 FT CDS /*tag=7a
 XX
 XX WO9904008-A2;
 XX
 XX 28-JAN-1999.
 XX
 XX 16-JUL-1996; 96WO-US14688.
 XX
 XX 27-JAN-1998; 98US-0014416.
 XX
 XX 18-JUL-1997; 97US-0053062.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Burk J, Emerson SU, Purcell RH, Yanagi M;
 XX
 XX WPI: 1999-132252/11.
 XX
 XX P-PSDB: AAM98022.
 XX
 XX New isolated hepatitis C virus nucleic acids - used to develop
 XX products for the diagnosis, prevention and treatment of HCV
 XX infections and for developing screening assays
 XX
 XX Claim 3; Fig 14A-F; 126pp; English.

The present sequence comprises the nucleic acid sequence of the
 genome of infectious hepatitis C virus (HCV) genotype 1b strain
 HC-J4 (ATCC 209596) that is capable of expressing this virus when
 transfected into cells. HC-J4 was obtained from acute phase plasma
 of a chimpanzee experimentally infected with serum containing
 HC-J4/91. The claimed infectious nucleic acid sequence can be used
 to produce chimeric genomes (see AAX24833) consisting of the open
 reading frames of infectious nucleic acid sequences of other
 genotypes (including genotypes 1-6) and subtypes (such as 1b, 2a,
 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also relates to
 the introduction of mutations or deletions into infectious nucleic
 acid sequences in order to produce an attenuated HCV virus suitable
 for vaccine development. Infectious nucleic acid sequences can
 also be used to produce attenuated virus via passage in vitro or in
 vivo of the viruses produced by transfection of a host cell with
 the infectious nucleic acid sequence. Vaccines comprising one or
 more polypeptides made from the infectious nucleic acid sequence are
 used to immunise mammals, especially humans, against hepatitis C.
 The nucleic acid sequences can also be used to induce protective
 immunity against the virus. The nucleic acid sequences or their
 encoded proteases (e.g. NS3 protease) can additionally be used to
 develop screening assays to identify antiviral agents for HCV.

Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;

Query Match 95.1%; Score 325.2; DB 20; Length 9595;
 Best Local Similarity 98.8%; Pred. No. 7.3e-87;
 Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagccccctgattgaggcgagacacccacatagatcctccctgtgaggaactactg 60
 DB 1 gccagccccctgattgaggcgagacacccacatagatcctccctgtgaggaactactg 60
 QY 61 tcttaacgagaagaagcgcttaacatgagcgtttagatgagtgctgacgctccagagcc 120
 DB 61 tcttaacgagaagaagcgcttaacatgagcgtttagatgagtgctgacgctccagagcc 120
 QY 121 ccccccctccgggagagacatagtgctgaggaacccgtgtagtaacccggaattgcag 180
 DB 121 ccccccctccgggagagacatagtgctgaggaacccgtgtagtaacccggaattgcag 180

QY 181 gacgaccgggtccttcttcttgatcaatccgcgtcaatgcttgagattggcgctgcc 240
DB 181 gacgaccgggtccttcttcttgatcaatccgcgtcaatgcttgagattggcgctgcc 239
QY 241 cgcgagactgctagcgcagtagtgtgtgtcgagaaaggccttggttactgctgtag 300
DB 240 cgcgagactgctagcgcagtagtgtgtgtcgagaaaggccttggttactgctgtag 299
QY 301 ggtgcttcgagtgccccggagaggtctcgttagaccgtgcacc 342
DB 300 ggtgcttcgagtgccccggagaggtctcgttagaccgtgcacc 341
RESULT 10
AAC86939 standard; DNA; 9595 BP.
ID AAC86939 standard; DNA; 9595 BP.
XX AAC86939;
AC AAC86939;
XX 02-APR-2001 (first entry)
DT 02-APR-2001 (first entry)
DE Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 1b.
XX Chimeric virus: bovine viral diarrhoea virus; BVDV; hepatitis C virus;
KM HCV; vaccine; viral inhibitor; antiviral; ss.
XX Hepatitis C virus.
XX Key Location/Qualifiers
FH CDS 342..9374
FT /*tag= a
XX WO200075352-A2.
PN 14-DEC-2000.
PD 02-JUN-2000; 2000MO-US15527.
PF 04-JUN-1999; 990S-0137817.
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Nam J, Bukh J, Emerson SU, Purcell RH;
PI WPI; 2001-071081/08.
DR P-PSDB; AAB31170.
XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus
PT genome in which the (non-)structural region has been replaced by
PT hepatitis C virus (HCV) genome useful for treating or preventing HCV
PT signs and symptoms
XX Disclosure; Fig 4A-F; 97pp; English.
XX The specification describes a nucleic acid comprising a chimeric virus
CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
CC the (non-)structural region has been replaced by the (non-)structural
CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
CC the chimeric virus and the chimeric virus are useful for identifying
CC cell lines capable of supporting the replication of these chimeric
CC viruses. In screening for neutralizing antibodies to HCV of different
CC genotypes, in the production of HCV-BVDV viruses, for the development
CC of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal,
CC in studying the molecular properties of HCV indirectly in vitro, and in
CC identifying inhibitors of viral enzyme activity which would be useful
CC as antiviral agents. Formulations or compositions comprising the
CC chimeric viruses may be used to treat or prevent the signs and symptoms
CC of HCV. The present sequence represents a HCV clone, which is used
CC to construct chimeric nucleic acids of the invention.
XX Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;

Query Match 95.18; Score 325.2; DB 22; Length 9595;
Best Local Similarity 98.88; Pred. No. 7.3e-87;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 ggcagccccctgattgggggagacacatccatgatgactccctctgaggaactactg 60
DB 1 ggcagccccctgattgggggagacacatccatgatgactccctctgaggaactactg 60
QY 61 tcttcacgacagaagagcgtctagcattgctatgagtgtcgtcaagctcagagcc 120
DB 61 tcttcacgacagaagagcgtctagcattgctatgagtgtcgtcaagctcagagcc 120
QY 121 ccccccctcccgaggagacatagtggtctcgagaaacgggtgagtaacacggaaattgccag 180
DB 121 ccccccctcccgaggagacatagtggtctcgagaaacgggtgagtaacacggaaattgccag 180
QY 181 gacgaccgggtccttcttcttgatcaatccgcgtcaatgcttgagattggcgctgcc 240
DB 181 gacgaccgggtccttcttcttgatcaatccgcgtcaatgcttgagattggcgctgcc 239
QY 241 cgcgagactgctagcgcagtagtgtgtgtcgagaaaggccttggttactgctgtag 300
DB 240 cgcgagactgctagcgcagtagtgtgtgtcgagaaaggccttggttactgctgtag 299
QY 301 ggtgcttcgagtgccccggagaggtctcgttagaccgtgcacc 342
DB 300 ggtgcttcgagtgccccggagaggtctcgttagaccgtgcacc 341
RESULT 11
AAF23492
ID AAF23492 standard; DNA; 9595 BP.
XX AAF23492;
AC AAF23492;
XX 21-MAR-2001 (first entry)
DT 21-MAR-2001 (first entry)
DE Infectious Hepatitis C virus 1b genotype.
XX GBV-B; hepatitis C virus; HCV; vaccine; ds.
OS Hepatitis C virus.
XX KC200075337-A1.
PN 14-DEC-2000.
PD 02-JUN-2000; 2000MO-US15293.
PF 04-JUN-1999; 990S-0137694.
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Bukh J, Yanagi M, Emerson SU, Purcell RH;
PI WPI; 2001-091214/10.
DR New infectious nucleic acids of the GB virus-B clone, useful for
PT indirectly studying the molecular properties of hepatitis C virus (HCV)
PT and in developing vaccines and therapeutics for HCV
XX Disclosure; Fig 7; 96pp; English.
XX The present invention relates to GB virus-B. The nucleic acid molecules
CC of the invention are useful for indirectly studying the molecular
CC properties of hepatitis C virus (HCV). The infectious nucleic acid
CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC in the development of vaccines and therapeutics for HCV.
XX Sequence 9595 BP; 1934 A; 2843 C; 2697 G; 2121 T; 0 other;

Query Match 95.18; Score 325.2; DB 22; Length 9595;

Best Local Similarity 98.8%; Pred. No. 7.3e-87;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagcccccgtatgggggagacacccacacatagatccctccctgtgaggaactacty 60
DB 1 gccagcccccgtatgggggagacacccacacatagatccctccctgtgaggaactacty 60
QY 61 tcttcacgagaagaacgctatgacatgagcttagatgagtgctgtcagacccctcagacc 120
DB 61 tcttcacgagaagaacgctatgacatgagcttagatgagtgctgtcagacccctcagacc 120
QY 121 ccccccctccgggagagacatagtgcttcgaggaaccggctgaaatagaccggaattccag 180
DB 121 ccccccctccgggagagacatagtgcttcgaggaaccggctgaaatagaccggaattccag 180
QY 181 gacgacggggtcccttctctgtgacaaatcccgctcaaatgtcctgtgagatttggtgctgccc 240
DB 181 gacgacggggtcccttctctgtgacaaatcccgctcaaatgtcctgtgagatttggtgctgccc 240
QY 241 cgcgagactgctagaccgaatagtgctgagctgcggaagaagcctgtgtactcctctatag 300
DB 241 cgcgagactgctagaccgaatagtgctgagctgcggaagaagcctgtgtactcctctatag 300
QY 301 ggtgctgtgagtgccccggaggtctcgtagaccgtgacc 342
DB 301 ggtgctgtgagtgccccggaggtctcgtagaccgtgacc 341

RESULT 12
AA024833
ID AA024833 standard; DNA; 9599 BP.
AC AA024833;
XX 21-JUN-1999 (first entry)
XX Infectious hepatitis C virus genotype 1a/1b chimera genome.
DE Infectious hepatitis C virus genotype 1a/1b chimera genome.
XX HCV; infectious clone; infection; diagnosis; therapy; vaccine;
KM screening; assay; antiviral; virucide; ss.
XX Hepatitis C virus.
OS
XX
XX Key Location/Qualifiers
FH 342..9377
FT CDS /*tag= a
XX
XX WO9904008-A2.
XX 28-JAN-1999.
XX
XX 16-JUL-1998; 98WO-US14688.
XX
XX 27-JAN-1998; 98US-0014416.
XX 18-JUL-1997; 97US-0053062.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Bukh J, Emerson SU, Purcell RH, Yanagi M;
XX
XX MPI: 1999-13252/11.
XX P-PSDB; AAM98021.
XX
XX New isolated hepatitis C virus nucleic acids - used to develop
XX products for the diagnosis, prevention and treatment of HCV
XX infections and for developing screening assays
XX
XX Clam 8; Fig 16A-F; 126pp; English.
XX
XX This nucleic acid sequence comprises the genome of infectious
XX hepatitis C virus (HCV) chimeric 1a/1b clone pH7C-J4 which
XX contains the nonstructural region of infectious genotype 1a strain
XX H77 (see AA024832) and the structural region of infectious genotype

1b strain HC-14 (see AA024833). The construction of such chimeric
nucleic acid sequences is expected to be of importance in studying
the growth and virulence properties of HCV and in the production
of HCV suitable for conferring protection against multiple
genotypes of HCV. The invention also relates to the introduction
of mutations or deletions into infectious nucleic acid sequences
in order to produce an attenuated HCV virus suitable for vaccine
development. Infectious nucleic acid sequences can also be used to
produce attenuated virus via passage in vitro or in vivo of the
viruses produced by transfection of a host cell with the infectious
nucleic acid sequence. Vaccines comprising one or more
polypeptides made from the infectious nucleic acid sequence are
used to immunise mammals, especially humans, against hepatitis C.
The nucleic acid sequences can also be used to induce protective
immunity against the virus. The nucleic acid sequences or their
encoded proteases (e.g. NS3 protease) can additionally be used to
develop screening assays to identify antiviral agents for HCV.

Sequence 9599 BP; 1889 A; 2873 C; 2724 G; 2113 T; 0 other:

Query Match 95.1%; Score 325.2; DB 20; Length 9599;

Best Local Similarity 98.8%; Pred. No. 7.3e-87;
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 gccagcccccgtatgggggagacacccacacatagatccctccctgtgaggaactacty 60
DB 1 gccagcccccgtatgggggagacacccacacatagatccctccctgtgaggaactacty 60
QY 61 tcttcacgagaagaacgctatgacatgagcttagatgagtgctgtcagacccctcagacc 120
DB 61 tcttcacgagaagaacgctatgacatgagcttagatgagtgctgtcagacccctcagacc 120
QY 121 ccccccctccgggagagacatagtgcttcgaggaaccggctgaaatagaccggaattccag 180
DB 121 ccccccctccgggagagacatagtgcttcgaggaaccggctgaaatagaccggaattccag 180
QY 181 gacgacggggtcccttctctgtgacaaatcccgctcaaatgtcctgtgagatttggtgctgccc 240
DB 181 gacgacggggtcccttctctgtgacaaatcccgctcaaatgtcctgtgagatttggtgctgccc 240
QY 241 cgcgagactgctagaccgaatagtgctgagctgcggaagaagcctgtgtactcctctatag 300
DB 241 cgcgagactgctagaccgaatagtgctgagctgcggaagaagcctgtgtactcctctatag 300
QY 301 ggtgctgtgagtgccccggaggtctcgtagaccgtgacc 342
DB 301 ggtgctgtgagtgccccggaggtctcgtagaccgtgacc 341

RESULT 13
AA032436
ID AA032436 standard; DNA; 7911 BP.
AC AA032436;
XX 26-APR-1993 (first entry)
XX
XX HCV antigen clone T7N1-30.
XX
XX Clon; Hepatitis C virus; HCV; core-envelope; NS1(gp70); NS2-NS4;
XX NS4-NS5; region; diagnostic method; antibody; suppress; control;
XX proteolytic; process; precursor; polypeptide; ss.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
FH 369..7898
FT CDS /*tag= a
XX
XX BP518313-A.
XX 16-DEC-1992.

```
XX PF 11-JUN-1992; 92EP-0109812.
XX XX 11-JUN-1991; 91JP-0139268.
XX PR 12-JUL-1991; 91JP-0172794.
XX PR 07-OCT-1991; 91JP-0287008.
XX PR 16-DEC-1991; 91JP-0323229.
XX PR 20-APR-1992; 92JP-0099957.
XX PA (MITU ) MITSUBISHI KASEI CORP.
XX PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
XX PI Teranishi Y;
XX DR WPI: 1992-417213/51.
XX DR P-PADB: AAR29527.
PT PT New hepatitis C virus gene and its encoded protein - used for
XX XX diagnosis and vaccinating against hepatitis C virus infections
PS PS Claim 1 and 3; Page 259-272; 305pp; English.
CC CC The sequences given in AA032436-539 are various clones which were all
CC derived from the isolated Hepatitis C Virus (HCV) gene of the
CC invention. AA032436 contains the entire protein gene, whereas AA032444-
CC 53 and AA032438-39, AA032454-68, AA032472-82 and AA032442, and
CC AA032502-12 and AA032443 encode the core-envelope, NS1(gp70), NS2-NS4
CC and NS4-NS5 regions respectively. AA032436 represents the entire gene
CC sequence. The HCV gene is useful in the development of a diagnostic
CC method which is more accurate and effective than conventional ones, in
CC the detection of antibodies raised against a wide range of HCVs which
CC have been hardly detected before. The complete gene may be used in an
CC in vitro screening system for a substance capable of specifically
CC suppressing or controlling a proteolytic processing of a precursor
CC polypeptide of HCV.
XX XX Sequence 7911 BP; 1556 A; 2363 C; 2292 G; 1700 T; 0 other:
SQ SQ
Query Match 95.0%; Score 324.8; DB 13; Length 7911;
Best Local Similarity 99.1%; Pred. No. 9,1e-87;
Matches 337; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 gccagccctcgtatggyggcgacactccaccatatcatcctccctgtgagaactacty 60
Db ttttgcgcgaagaacggtctaggccagttagatagtgctgtgacgtccagcgc 120
QY 61 ttcttcgcgagaagaacggtctaggccagttagatagtgctgtgacgtccagcgc 120
Db 88 tccttaacgcaaaaacggtctcctagccatggcgttagatagtgctgtgacgtccagcgc 147
QY 121 ccccccctccggaagacgatagtgctcgcgaaacgggtgatcacccggaattgccag 180
Db 148 cccccctcccggaagacgatagtgctcgcgaaacgggtgatcacccggaattgccag 207
QY 181 gacgacgcgggtccttctcttgatataatcccgctcaaatgacctggagattggcggtgcc 240
Db 208 gacgacgcgggtccttctcttgatataatcccgctcaaatgacctggagattggcggtgcc 266
QY 241 ccgcgagactcgtacgagcgatggtgtgtggtcgcgaaaggcctgtggtactccctgatat 300
Db 267 ccgcgagactcgtacgagcgatggtgtgtggtcgcgaaaggcctgtggtactccctgatat 326
QY 301 ggtgtctgcgagtgcccggaagagctcgtltagaccgtgtca 340
Db 327 ggtgtctgcgagtgcccggaagagctcgtltagaccgtgtca 366
AA062945:
ID ID AA062945 standard; DNA: 341 BP.
XX XX
```

DT		19-SEP-1994	(first entry)
XX	XX	5'	untranslated region of HCV.
KW	XX	Hepatitis C virus;	translation; control; pestivirus; ss.
CS	XX	Hepatitis C virus.	
RH	XX	key	location/qualifiers
FT	XX	misc-structure	1..23
FT	XX	/tag= "a"	
FT	XX	/note= "5' hairpin structure"	
FE	XX	misc-feature	291..318
FT	XX	/tag= b	
FT	XX	/note= "pestivirus homology box IV"	
FT	XX	misc-feature,	277..300
FT	XX	/tag= c	
FT	XX	/note= "AS5 region"	
FN	XX	MO8408002-A.	
PD	XX	16-Apr-1994.	
PF	XX	28-SEP-1993;	93WO-US09200.
PR	XX	28-SEP-1992;	9ZUS-0952799.
XX	XX	(CHIR) CHIRON CORP.	
BA	XX	Hang JH,	Houghton M, Selby MJ, Spaete RR, Suh BS;
PI	XX	YDQ BJ;	
DR	XX	WFI: 1994-135569/16.	
XX	XX	Controlling translation of viral peptide from viral nucleic acid	
PT	XX	- using sense and antisense sequences from 5'-untranslated region	
BT	XX	of pestivirus and hepatitis C virus	
PS	XX	Disclosure: Page 31; 38pp; English.	
CC	XX	The sequence is that of the 5' untranslated region of the Hepatitis	
CC	XX	C virus. The 5' hairpin pref. forms a triple helix with sequences	
CC	XX	associated with the hairpin, inhibiting cleavage from the remaining	
CC	XX	portion of the mRNA. A cleavage area of 50 nucleotides spanning	
CC	XX	position 145 is capable of binding to an area defined by the site at	
CC	XX	which long full-length genomic RNA is cleaved to form shorter,	
CC	XX	singnomic RNA. The AS5 region is capable of binding to a region	
CC	XX	that overlaps the pestivirus homology box IV. The sequence can be	
CC	XX	used for controlling the translation of viral peptides from HCV	
CC	XX	nucleic acid, with particular applications to pestivirus and HCV.	
CC	XX	See also AA062946.	
SO	XX	Sequence 341 BP; 64 A; 103 C; 105 G; 69 T; 0 other;	
Query Match		94.6%; Score 323.6; DB 15; Length 341;	
Best Local Similarity		98.5%; Pred. No. 9.4e-87;	
Matches 337; Conservative		0; Mismatches 4; Indels 1; Gaps	
Dp		1 gccagcccccctgattggaggcgacactcaccatagatactccctgttgagaactacty 60	
		1 gccagccccctgatggggcgacctccacatgatactcctccctgtgagaactacty 60	
Gy		tcttaacgcgaagaaacgctctagcatgycgttatgatatgagtgtgtgcagcctcaagcc 120	
Fb		51 tcttaacgcgaagaaacgctctagcatgycgttatgatatgagtgtgtgcagcctcaagac 120	
Gy		121 cccccctccggagagacatagtgcttcgcggaacccgltgagttacaccggaatgccag 180	
Fb		121 cccccctccggagagacatagtgcttcgcggaacccgltgagttacaccggaatgccag 180	
Gy		141 gacgacccggtctcttcttgatcataccccgcgtcaatgctctlgagatttggcgtgcccc 240	

```

DB 181 gagcagcgggtctctcttcgataca-ccgcctcaatgctcgtgagattcggcgctgcccc 239
OY 241 cgcgagactgctagaccgagtaagtgtggtgcgaaaggccttgtagctactgctagatag 300
DB 240 cgcgagactgctagaccgagtagtgtggtgcgaaaggccttgtagctactgctagatag 239
OY 301 ggtgctgctgagtgcccgaggaggtctcgtatagaccgtgacc 342
DB 300 ggtgctgctgagtgcccgaggaggtctcgtatagaccgtgacc 341

RESULT 15
AAZ08979 standard; DNA; 341 BP.
AC AAZ08979;
DE 02-NOV-1999 (first entry)
DE Hepatitis C virus 5' UTR used as promoter for a HCV genetic vaccine.
KW Genetic vaccine; immunisation; humoral response; cellular response;
KW non-structural protein; NS protein; antigen; ds.
OS Hepatitis C virus.
XX WO9338880-A1.
XX 05-AUG-1999.
XX 28-JAN-1999; 99WO-US01823.
XX 30-JAN-1998; 98US-0073156.
PA (GENO ) GEN HOSPITAL CORP.
PI Encke J, Wands J;
XX WPI: 1999-494077/41.
DR New hepatitis virus nucleic acids for, e.g. inducing an immune
PT response against the virus
PS Claim 6: Page 10-11; 41pp; English.

This sequence represents the hepatitis C virus (HCV) 5' untranslated
region (5' UTR). This is preferred for use as a promoter in a
recombinant DNA vaccine against hepatitis C. The genetic
vaccine comprises regulatory elements capable of directing
expression in humans and one or more HCV non-structural
(NS) proteins. The HCV genome encodes three NS proteins: NS3, a serine
protease/viral helicase; NS4; and NS5, the viral RNA-dependent RNA
polymerase. In HCV infection, these and other virally encoded proteins
are produced by proteolytic processing of a precursor polypeptide.
However, for use in a genetic vaccine, the sequences coding for NS
proteins had to be engineered so that each NS sequence had its own
initiation and stop codons. Restriction sites were also engineered
into the gene fragments to aid subcloning. The gene fragments were
isolated and mutated using PCR. The NS3 coding sequence was modified
using PCR primers AAZ08980 and AAZ08981, the NS4 coding sequence
modified using AAZ08984 and AAZ08985, and the NS5 coding sequence
modified using AAZ08986 and AAZ08987. Studies in mice immunised with
constructs expressing one NS protein were found to elicit strong
antigen-specific immune responses in both arms of the immune system. This
demonstrated that the NS proteins are better antigens for stimulating
humoral immune responses as compared with previous studies using the HCV
core structural protein. The genetic vaccine is useful for inducing
an immune response (cellular or humoral) against hepatitis C virus in
a human uninfected by the virus, and for immunising a human susceptible
to hepatitis C viral infection by inducing an immune response. The
composition is also useful for treating a human infected with hepatitis
C virus, by induction of an immune response. The advantage of this

```

```

CC method of immunisation compared with immunisations with soluble
CC recombinant proteins or peptides, is the ability to induce a strong
CC inflammatory CD4+ T cell response as well as cytotoxic T cell activity.
CC In addition, the new recombinant genetic vaccine is more suitable for
CC immunisation; unlike synthetic peptides which only have a limited
CC number of epitopes available for stimulation of the host response.
SQ Sequence 341 BP; 63 A; 103 C; 106 G; 69 T; 0 other;

Query/Match          94.6%; Score 323.6; DB 20; Length 341;
Best Local Similarity 98.5%; Pred No. 9.4e-87;
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

OY 1 gccagcccccctgattggggcgacactccacacatagatcccccgtgagaactactg 60
DB 1 gccagcccccgtattggggcgacactccacacatagatcccccgtgagaactactg 60
OY 61 tcttcacgagaagaagcgtctagcgaatggttagttagtgcgtgcagcctcagacc 120
DB 61 tcttcacgagaagaagcgtctagcgaatggttagttagtgcgtgcagcctcagacc 120
OY 121 ccccccctccggagagacatagtgctgcgaaacggtgagtaaccggaattgccag 180
DB 121 ccccccctccggagagacatagtgctgcgaaacggtgagtaaccggaattgccag 180
OY 181 gagcagcgggtcccttcttgatcaatcccgctcaatgctctgagaatttggtgcgtgcc 240
DB 181 gagcagcgggtcccttcttgatcaatcccgctcaatgctctgagaatttggtgcgtgcc 239
OY 241 cgcgagactgctagaccgagtagtgtggtgcgaaaggccttgtagctactgctatag 300
DB 240 cgcgagactgctagaccgagtagtgtggtgcgaaaggccttgtagctactgctatag 299
OY 301 ggtgctgctgagtgcccgaggaggtctcgtatagaccgtgacc 342
DB 300 ggtgctgctgagtgcccgaggaggtctcgtatagaccgtgacc 341

```

Search completed: March 1, 2002, 15:30:43
Job time: 25051 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 11:35:41 ; Search time 8196.34 Seconds

(without alignments)
448,378 Million cell updates/sec

Title: US-09-763-836-7

Perfect score: 342
Sequence: 1 gccagcccccgtatg99g9c.....ggtctcgtagacgltgaccc 342

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gp_est1:*
11: gp_est2:*
12: gp_hic:*
13: gp_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40.8	11.9	925	13	CNS0091P
2	38	11.1	925	13	CNS0091P
3	35.6	10.4	844	13	CNS0052P
4	35.2	10.3	600	13	AC038955
5	34.2	10.0	548	10	AW958779
6	34.2	10.0	610	10	AW654578
7	34	9.9	394	11	BF901361
8	34	9.9	440	11	BE845986
9	33.4	9.8	421	10	AV397071
10	33.4	9.8	432	10	AA196987
11	33.4	9.8	777	11	BF859305
12	32.8	9.6	371	11	BF508846

13	32.8	9.6	405	13	AO815897
14	32.8	9.6	849	11	BE759010
15	32.8	9.6	927	11	BE492855
16	32.8	9.6	1101	13	CNS005R8
17	32.6	9.5	645	13	CNS01213
18	32.6	9.5	952	10	AL563471
19	32.4	9.5	339	10	AV686380
20	32.4	9.5	369	10	AV693842
21	32.4	9.5	370	10	AV693850
22	32.4	9.5	575	11	BF001581
23	32.4	9.5	577	13	AA224104
24	32.4	9.5	961	13	CNS01087
25	32.2	9.4	400	10	BE498161
26	32.2	9.4	470	11	BE263335
27	32.2	9.4	977	13	CNS00JX7
28	32.2	9.4	430	11	HS7082
29	32	9.4	897	11	BE031227
30	31.8	9.3	436	10	AI094710
31	31.8	9.3	778	10	BE131094
32	31.6	9.2	530	11	BE349419
33	31.6	9.2	793	10	BE408210
34	31.4	9.2	638	13	AZ003859
35	31.4	9.2	901	11	BI414151
36	31.4	9.2	949	11	BE799450
37	31.2	9.1	442	10	BE728496
38	31.2	9.1	549	10	BE424949
39	31.2	9.1	844	13	CNS0052P
40	31.2	9.1	1003	10	AL523533
41	31.2	9.1	1006	11	BI104877
42	31	9.1	387	10	AA921613
43	31	9.1	647	10	AV650450
44	31	9.1	648	10	AV650417
45	31	9.1	683	13	AZ125374

ALIGNMENTS

RESULT	1
LOCUS	CNS0091P
DEFINITION	Drosophila melanogaster genome survey sequence TERT end of BAC # BACR19D16 of RPCT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL053013
VERSION	AL053013.1 GI:4934461
KEYWORDS	GSS.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS	1 (bases 1 to 925)
TITLE	Genoscope.
COMMENT	Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila genome project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .

```

FEATURES                                     Location/Qualifiers
source                                         1..925
                                                /organism="Drosophila melanogaster"
                                                /db_xref="taxon:7227"
                                                /clone_11b="Rpci-98"
                                                /clone="BACR19D16"
                                                /note="end : TTT3"

BASE COUNT      120 a       61 c       61 g       172 t     511 others
ORIGIN

Query Match          11.9%; Score 40.8; DB 13; Length 925;
Best Local Similarity 14.9%; Pred. No. 0.28;
Matches    43; Conservative 117; Mismatches 129; Indels   0; Gaps   0;

Qy  48 tgaggaactacgtctccgcgagaaggctatgaccatggcgttagatagtgtcg 107
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  502 TTATTTANNNNNANNNANNNANNNNNAGCCSMCGKKGCTTBGTGTTTTSSGGTG 561

Qy  108 cagcctcagagcccccccctcccgagagacatagtgtctcgaaacctgtagtaa 167
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  562 KCSSSGSBSCACSSCSGCCSCSSSCBCCCCSCSYCCSSBSBKSTSBSCGCC 621

Qy  168 ccggaatccgagagcagcagcggtctcttcttgatcaatccgcgtcaatgcctgagat 227
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  622 SSKSVCGTSCSSSSCSCSSSSTSSTSSSTTSKSSSSSSSSSYTTSKSASGSG 681

Qy  228 ttgggcgltgcccccgcgaactcgtctaagcgaatagtgttggttcgcgaagacctgttg 287
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  682 SMSAGGSGSGTGSTSSSSSSSTSSSSSVSGSKMSPTBSGSBSBGSSSSSSTBSBS 741

Qy  288 tactcgcgatattggtgtcgttgtcgagtgtcccgagaggtctgttaacgg 336
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  742 CTSTSSSSSSYSTSCCTCCCTCCSYSSSYSSSTSSSSTSWGSTSSSSS 790

RESULT 2
LOCUS   CNS0091P/c
DEFINITION Drosophila melanogaster genome survey sequence TERT3 end of BAC #
            BACK19D16 of Rpci-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION AL053013
VERSION   AL053013.1 GI:4934461
KEYWORDS
SOURCE    GSS:
           fruit fly.
           Drosophila melanogaster
           Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
           Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
           Muscomorpha; Ephyndroidea; Drosophilidae; Drosophila.
REFERENCE Genoscope.
AUTHORS   Submitted Submission
TITLE      Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
JOURNAL   BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT    - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osogawa and
            Aaron Mamoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            pl and Est libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES                                     Location/Qualifiers
source                                         1..925

```

/organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="RPc1-98"
 /clone="BAC119D16"
 /note="end : TET3"

BASE COLUMN 120 a 61 c 61 g 172 t 511 others
 ORIGIN

Query Match 11.1%; Score 38; DB 13; Length 925;
 Best Local Similarity 13.7%; Pred. No. 1.7; Mismatches 104; Indels 0; Gaps 0;
 Matches 53; Conservative 117;

QY 69 cagaagcgtctagccatgccttagtatgaatgtgtgtcagacccagcccccctc 128
 B 829 CGSASARGVKVRASGAGKRGSGGASASHSSSSSACBSSSSSCASCSMASSSSSASS 770
 QY 129 ccggagagacatagtgtgtctcgcgaagccggtgtgatacacgcggaattgccaggaacgc 188
 B 769 RSHSGCGAGGSASSSSSSSSSSSAGAGVSASSSSSSSSSVSCSVASMSGCSBBS 710
 QY 189 ggtcctcttgatcaatcccgccaatgagattgtggcgtgcgcgcgcgaagc 248
 B 709 SSSASASSSSSSSSSASCACSCCCTCMTWSCCCTSMASMSARSSSSSSSSSSSMASSS 650
 QY 249 tctctagcgaagtgtgtgtgtgtgcgaagccctgtgttactgcctataggtgtctg 308
 B 649 AASSSSASSSSSSSSSSSACGBSMSSGCGGGSVASSGMSSSSVSSSGGRSGSGGGG 590
 QY 309 cgatgtcccgaggagg 324
 B 589 VCGSSGSSSSGSSSGC 574

RESULT 3
 CNS0052P
 LOCUS
 DEFINITION
 Drosophila melanogaster genome survey sequence TET3 end of BAC #
 BAC119D16 of RPc1-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL056652
 AL056652.1 GI:4932342
 GSS.

ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 844)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osogawa and
 Aaron Mamoser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPc1-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 Location/Qualifiers
 1..844

FEATURES
 SOURCE
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="BAC119D16"

BASE COUNT	261 a	112 c	92 g	35 t	344 others
ORIGIN					
Query Match	10.4%	Score 35.6;	DB 13;	Length 844;	
Best Local Similarity	15.7%	Pred. No. 8;			
Matches 37;	Conservative 106;	Mismatches 92;	Indels 0;	Gaps 0;	
OY	104	ctgtgacgcttcacagccccccctccctccgggagagacatagtggtctgtggagacgggtgag	163		
Db	403	SRSRMWSGCCSACSSGCVSGSSGAMSSSCCCVSVSGMAMSSSCCGMSASSSSSGG	462		
OY	164	tacacgggaattgcccaagacacgcggctcttcttgatcaatccgcctcaatgccttg	223		
Db	463	VSSVSGGAAVAGRGCMCCSMVCCCMCCSMCCSVSVCAVCGSVSGSVRAVAVGCSVGG	522		
OY	224	agatttggcgtgcgtcccccgcgagactgctgaagcgaagtgtgtgtgtgcgcgaagacctt	283		
Db	523	RVGGSSRRGRAGSSRGGRGSSVSSGVSVSSSSSVGMCACASASVSCSBSVSAASSGGB	582		
OY	284	gtgtgacgtccctgataaggtgtctgtgcggtgtgccccgggagagtgctgtgtagccgtg	338		
Db	583	VSRRGCGRCVGGGAGVGGSRVSSGSSSSSSSSGCGSGVSRSRSGANGVRVGGG	637		
RESULT 4					
LOCUS	A0538955	600 bp	DNA	GSS	19-MAY-1999
DEFINITION	RPCT-11-343B22.TF RPCT-11 Homo sapiens genomic clone RPCT-11-343B22				
ACCESSION	A0538955				
VERSION	A0538955.1 GI:4868938				
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and Venter,J.C.				
TITLE	Use of BAC End Sequences from Library RPCT-11 for Sequence-Ready Map Building				
JOURNAL	Unpublished (1997)				
COMMENT	other.GSS: RPCT-11-343B22.TV Contact: Shaying Zhao, William Niernan, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are derived from the human BAC library RPCT-11. For BAC library availability, please contact Pieter de Jong (pieter@edj.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html. Seq primer: SP6 Class: BAC ends.				
FEATURES	Location/Qualifiers				
source	1..600				
	/organism="Homo sapiens"				
	/db_xref="GDB:7631375"				
	/db_xref="taxon:9606"				
	/clone="RPCT-11-343B22"				
	/clone_lib="RPCT-11"				
	/sex="Male"				
	/cell_type="lymphocytes"				
	/note="Vector: pbac3.6; Site_1: EcoRI; Site_2: EcoRI; RPCT11 Human Male BAC Library"				

BASE COUNT	129 a	155 c	151 g	165 t
ORIGIN				
Query_Match		10.3%	Score 35.2;	DB 13; Length 600;
Best Local Similarity		53.7%;	Pred. No. 9.5;	
Matches	73; Conservative	0;	Mismatches	63; Indels 0; Gaps 0;
Cy	113 gagagccatagtgctgcgcgaacccgcygagaaaccggaattgcacagagccggtc	192		
Db	332 GAGGAATAATGATGTATCCGAGACCACTGAGAGCTTAACTTATGCTTGAGGGCTTGCA	323		
Cy	193 ctctcttgatcaatcccgctcaatgctctggagatttggcgcgcgcgcgcgcgcgcgcgc	252		
Db	332 CTGTCAGAGAAACCAACCAAGATCAATTTGTGGATATATGTTTCCACCCCTCAGGAACTGCT	263		
Cy	253 agccgagatgagtgtg 268			
Db	252 AGCCCTACTGGGGTGG 247			
RESULT 1 5				
LOCUS	AM958779	548 bp	mRNA	EST
DEFINITION	EST370849	MAGE	resesquences,	MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION	AM958779			
VERSION	AM958779.1	GI:8148463		
KEYWORDS	EST.			
SOURCE	human;			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 548)			
COMMENT	Hedge, P., Ol, R., Abernathy, K., Daarp, S., Gaspar, R., Gay, C., Holt			
JOURNAL	, I.E., Seed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and			
QUACKENBUSH, J.	Quackenbush, J.			
ASSESSMENT	Assessment of gene expression patterns in a model of colon tumor			
metastasis	metastasis using a 19,200 element cDNA microarray			
UNPUBLISHED	Unpublished (2000)			
CONTACT	Contact: John Quackenbush			
INSTITUTE	The Institute for Genomic Research			
ADDRESS	9712 Medical Center Dr., Rockville, MD 20850, USA			
TEL	Tel: 301 838 3528			
FAX	Fax: 301 838 0208			
EMAIL	Email: john@tigr.org			
PI	PI: 125			
Seq primer:	Reverse.			
Location/Qualifiers	1..548			
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/clone_lib="MAGE resesquences, MAGE"				
/note="Vector: pBluescriptSkm"				
BASE COUNT	111 a	161 c	149 g	127 t
ORIGIN				
Query_Match		10.0%	Score 34.2;	DB 10; Length 548;
Best Local Similarity		50.3%;	Pred. No. 18;	
Matches	84; Conservative	0;	Mismatches	83; Indels 0; Gaps 0;
Cy	8 cccatattgagggagacactccac	67		
Db	155 CTCACATATTGATGGGCTCTGAGAGCCATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	214		
Cy	68 gacgaagacgctac	127		
Db	215 CGTGCGCTGCTTGATGATATGATATGCGCTACGTGGCGTGCAGCATTTGGGATTCAGGCT	274		
Cy	128 cccgagagac	174		
Db	275 ACCTAGAGGGGATCGGGCCAGGAAACCTTGATTCAGCAAGCAAT 321			

RESULT	6	AV654578	610 bp	MRNA	EST	07-SEP-2000
LOCUS	AV654578	GLC Homo sapiens	cdna	GLC	103	3' mRNA sequence.
DEFINITION	AV654578	GLC Homo sapiens	cdna	GLC	103	3' mRNA sequence.
ACCESSION	AV654578	GLC Homo sapiens	cdna	GLC	103	3' mRNA sequence.
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
AUTHORS	1 (bases 1 to 610)					
	Qian,B., Wu,T., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z., and Han,Z.					
TITLE	Homo sapiens CDNA clone					
JOURNAL	Unpublished (2000)					
COMMENT	Contact: Zeguan Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203 P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.					
FEATURES	Location/Qualifiers					
SOURCE	1..610					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone="GLC10"					
	/clone_id="GLC"					
	/tissue_type="corresponding non cancerous liver tissue"					
	/dev_stage="Adult"					
	/lab_host="SOLR"					
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"					
BASE COUNT	118 a 186 c 184 g 121 t					1 others
ORIGIN						
Query Match	10.0%; Score 34.2; DB 10; Length 610;					
Best Local Similarity	50.3%; Pred. NO. 18;					
Matches	84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;					
QY	8 cccctagtgaggcgacatccacccatagatcctccctgtgaggaactactgtcttcac					67
DB	317 CTCACATATTGAGGGCTCGAAGCCCTAGTTCGTCTCTCTGAGAGAGCTGTGGCTTCCC					376
QY	68 gcaagaagcgtctagatcgcgtgctagtagtgcgtgcagcctccagccccccct					127
DB	377 CCGTGGCTGTTGGTGCATGCATTTAGCGCTAGCTGGCGTGCAGCATTTGGGATTCACAGGCT					436
QY	128 cccggagagccatagtggtctgcggagccggtgagttacaccggaat					174
DB	437 ACCCTAGAGGGGCGCATCGGCCAGGAAACCTCGGATTACAGCAAT					483
RESULT	7	BF901361	394 bp	MRNA	EST	18-JAN-2001
LOCUS	BF901361	394 bp	MRNA	EST	18-JAN-2001	
DEFINITION	PM2-MT0199-081200-001-a10 MT0199 Homo sapiens					CDNA, mRNA sequence.
ACCESSION	BF901361					
VERSION	BF901361.1					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
REFERENCE	1 (bases 1 to 394)					
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,					

FEATURES	Source
Query Match	9.9%; Score 34; DB 11; Length 394;
Best Local Similarity	50.6%; Pred. NO. 18;
Matches	82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 120	ccccccctccggagagacatagtggtctcgagacgggtgagtacacggaaattgcc 179
DB 177	CCTCGCGACGCCCTCGCGCCGCGAGGACGACGCGGCCCGCAGAAATACACGCGCCAGGCGMA 118
QY 180	gagacgacgggtgccttcttgatcaatcccgctcaatgacctgaggaattggcgctgcc 239
DB 117	TGAGCACTGTACGACACAGGTTCTCCATTCACGATCCCTCGACGACGCCCGGCTTCACCG 58
QY 210	ccgcgagactgctagcgcagtagtgggtggtcgcaaaagcc 281
DB 57	TAGAGCACTGCATCGCTCTGCGCCTGGCCTGGACAGGAC 16
RESULT	8
LOCUS	BE845986 440 bp mRNA EST 25-SEP-2000
DEFINITION	232495 BANC 5BOV Bos taurus CDNA 5', mRNA sequence.
ACCESSION	BE845986
VERSION	BE845986.1 GI:10282801
KEYWORDS	EST.
SOURCE	COW.
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos. 1 (bases 1 to 440)
AUTHORS	Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and Wells,K.D.
TITLE	Mapping of Expressed Sequence Tags from a normalized bovine mammary

gland cDNA library
unpublished (2000)
Contact: Sonstegard TS

Email: tdadelpsi.barc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
V0.9809094.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.

1. .440

```

/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 580V"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPOR6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues." at eight physiological, developmental, and disease
states."

```

90 a	135 c	133 g	82 t
------	-------	-------	------

Query Match	9.9%	Score 34;	DB 11;	Length 440;
Best Local Similarity	50.6%;	Pred. No. 19;		
Matches 82;	Conservative 0;	Mismatches 80;	Indels 0;	Gaps 0;

OY 59 tgccttcacgcagaanaagcgccttagccatgcgctagtatggtgtcgtcgaagccctccagg 118
 || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 TGACTTTAAGCAGCAACCTCCCGTCCACCACGCAgTTCCGTGACTTCAGCCGCCCATGGG 248

qy 119 cccccccctccgagagacatagtcgtcgcggaaccgylgatlacaccggaattgcc 178
 || |||| | | | | | | | | | | | | | | | |
 Db 249 CCACCCCTCGATGCTGCAGATTACCTTATTTCCCTGGAAAGGCTCATTAAGGCGACACTC 308

```

QY      179 aggacgacccggtccttctcttgatcaatcccgctcaatgcc 220
          || | ||||| | | | | | | | | | | | | | |
Db      309 AGCCCAcGCGGGTCCTCTTCTGGCTCCTCCGCTGATCCC 350

```

LOCUS	421 bp	mRNA	EST	29-SEP-2000
DEFINITION	AV397071	Chlamydomonas reinhardtii C9	Chlamydomonas reinhardtii	
	CDNA clone CL75b06_r	mRNA sequence.		

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 421)
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green
alga, *Chlamydomonas reinhardtii*. I. Generation of 343
non-redundant expressed sequence tags
DNA Res. 6 (6), 369-373 (1999)

Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: yanakamuekazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.
Location/Qualifiers

Location/Qualifiers

```
1, .421
/organism="Chlamydomonas reinhardtii"
```

```

/note=Vector: pBluescript1 SK-, site_1: EcoRI, site_2:
XhoI"
BASE COUNT      83 a      147 c      114 g      77 t
ORIGIN          :

```

83 a	147 c	114 g	77 t
------	-------	-------	------

```

Query: Match 9.8%; Score 33.4; DB 10, Length 421;
      Best Local Similarity 55.7%; Pred. No. 28;
      Matches 64, Conservative 0; Mismatches 51; Indels 0; Gaps 0;

```

DQ 163 gtacaccggaattgcagagcagccggcttccttcttgatcaalcccgcctcatgctgt 222
+ + + + +
P 1 + + + + +
Db 116 GCACTCAGCACCAAGCAGGTGTTGCCGGCTGTTATTGGGTTGAGCCAGCGAGCGCGTG 57

Dh
66 GATGCTGGGACCGAGACGCCTGTGGCCCTAGCTGTAAACCGACCAGAGA 2

DEFINITION
zq60h12.r1 Striatagene neuroepithelium (#337231) Homo sapiens CDNA
clone IMAGE:546055 5' similar to gb:J03040 SPAC PRECURSOR (HUMAN)
); mRNA sequence.

1 GI:1792578

a; n

S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins

TITLE. Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)

on University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through ILTIL; contact the
IMAGE Consortium (info@image.ill.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 227.

1. .432

```

/organism="Homo sapiens"
/db_xref="GDB:5216306"
/db_xref="taxon:9606"
/clone IMAGE:646055"
/clone_lib="Stratagene neuroepithelium (#937231)"
/dev_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: plasmid SK-; Site:1: EcoRI; Site:2:
XhoI; cloned unidirectionally. Primer: Oligo dT, NT2
cells (Ntera-2/ci.D1) Induced with Retinoic Acid for 24
hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5
adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor

```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

Run on: March 1, 2002, 15:26:13 ; Search time 115.48 Seconds
(without alignments)
670.726 Million cell updates/sec

Title: US-09-763-836-7

Sequence: 1 gccagccccctgatggggc.....ggtctcgtagaccgtgcacc 342

Scoring table: IDENTITY_NUC

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	08
Minimum Match	08	1000

Listing first 45 summaries

Database : Issued_Patents_NA:*

- ```

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID                | Description        |  |
|------------|-------|-------------|--------|----|-------------------|--------------------|--|
|            |       |             |        |    |                   |                    |  |
| 1          | 325.2 | 95.1        | 341    | 3  | US-09-014-416-48  | Sequence 48, Appl  |  |
| 2          | 325.2 | 95.1        | 9401   | 2  | US-08-432-693-1   | Sequence 1, Appl1  |  |
| 3          | 325.2 | 95.1        | 9416   | 3  | US-08-811-566-19  | Sequence 19, Appl  |  |
| 4          | 325.2 | 95.1        | 9595   | 3  | US-09-014-416-4   | Sequence 4, Appl1  |  |
| 5          | 325.2 | 95.1        | 9599   | 3  | US-09-014-416-6   | Sequence 6, Appl1  |  |
| 6          | 323.6 | 94.6        | 341    | 2  | US-08-440-209-1   | Sequence 1, Appl1  |  |
| 7          | 323.6 | 94.6        | 341    | 3  | US-08-854-531-4   | Sequence 4, Appl1  |  |
| 8          | 323.6 | 94.6        | 341    | 3  | US-08-439-996-1   | Sequence 1, Appl1  |  |
| 9          | 323.6 | 94.6        | 341    | 3  | US-09-014-416-47  | Sequence 47, Appl  |  |
| 10         | 323.6 | 94.6        | 341    | 4  | US-08-869-380-4   | Sequence 4, Appl1  |  |
| 11         | 323.6 | 94.6        | 341    | 5  | PCT-US95-13552-4  | Sequence 4, Appl1  |  |
| 12         | 323.6 | 94.6        | 342    | 3  | US-08-474-7008-39 | Sequence 39, Appl1 |  |
| 13         | 323.6 | 94.6        | 350    | 2  | US-07-863-622-1   | Sequence 1, Appl1  |  |
| 14         | 323.6 | 94.6        | 350    | 5  | PCT-US93-03266-1  | Sequence 1, Appl1  |  |
| 15         | 323.6 | 94.6        | 686    | 4  | US-08-988-3218-37 | Sequence 37, Appl  |  |
| 16         | 323.6 | 94.6        | 686    | 4  | US-08-397-2208-25 | Sequence 25, Appl  |  |
| 17         | 323.6 | 94.6        | 780    | 3  | US-08-474-7008-45 | Sequence 45, Appl  |  |
| 18         | 323.6 | 94.6        | 923    | 4  | US-08-869-380-1   | Sequence 1, Appl1  |  |
| 19         | 323.6 | 94.6        | 923    | 4  | PCT-US95-13552-14 | Sequence 14, Appl  |  |
| 20         | 323.6 | 94.6        | 9401   | 1  | US-07-910-760-9   | Sequence 9, Appl1  |  |
| 21         | 323.6 | 94.6        | 9401   | 1  | US-08-440-519-9   | Sequence 9, Appl1  |  |
| 22         | 323.6 | 94.6        | 9401   | 5  | PCT-US91-02225-9  | Sequence 9, Appl1  |  |
| 23         | 322   | 94.2        | 341    | 3  | US-09-014-416-19  | Sequence 49, Appl  |  |
| 24         | 322   | 94.2        | 9599   | 3  | US-09-014-416-2   | Sequence 2, Appl1  |  |
| 25         | 322   | 94.2        | 9646   | 3  | US-08-811-566-1   | Sequence 1, Appl1  |  |
| 26         | 322   | 94.2        | 12980  | 3  | US-08-811-566-5   | Sequence 5, Appl1  |  |
| 27         | 315.4 | 92.2        | 1499   | 1  | US-08-324-977-3   | Sequence 3, Appl1  |  |

|    |       |      |      |   |                   |                    |
|----|-------|------|------|---|-------------------|--------------------|
| 28 | 315.4 | 92.2 | 1499 | 2 | US-08-384-616-3   | Sequence 3, Appl1  |
| 29 | 315.4 | 92.2 | 1499 | 2 | US-08-384-668A-3  | Sequence 3, Appl1  |
| 30 | 315.4 | 92.2 | 1499 | 4 | US-09-315-850-3   | Sequence 3, Appl1  |
| 31 | 315.4 | 92.2 | 1496 | 1 | US-08-324-977-1   | Sequence 1, Appl1  |
| 32 | 315.4 | 92.2 | 9416 | 2 | US-08-384-616-1   | Sequence 1, Appl1  |
| 33 | 315.4 | 92.2 | 9416 | 2 | US-08-904-688A-1  | Sequence 1, Appl1  |
| 34 | 315.4 | 92.2 | 9416 | 2 | US-09-315-850-1   | Sequence 1, Appl1  |
| 35 | 311.4 | 91.1 | 334  | 2 | US-08-470-426B-15 | Sequence 15, Appl1 |
| 36 | 311.4 | 91.1 | 1863 | 2 | US-08-470-426B-14 | Sequence 14, Appl1 |
| 37 | 306.6 | 89.6 | 334  | 2 | US-08-470-426B-13 | Sequence 13, Appl1 |
| 38 | 306.6 | 89.6 | 1863 | 2 | US-08-470-426B-13 | Sequence 13, Appl1 |
| 39 | 301.6 | 88.2 | 9185 | 3 | US-08-444-818-12  | Sequence 12, Appl1 |
| 40 | 301.6 | 88.2 | 9185 | 3 | US-08-444-818-13  | Sequence 12, Appl1 |
| 41 | 301.6 | 88.2 | 9379 | 3 | US-08-444-818-11  | Sequence 12, Appl1 |
| 42 | 301.6 | 88.2 | 9379 | 4 | US-09-388-874-1   | Sequence 1, Appl1  |
| 43 | 293.6 | 85.8 | 9589 | 1 | US-07-925-695-1   | Sequence 1, Appl1  |
| 44 | 293.6 | 85.8 | 9589 | 1 | US-07-925-695-2   | Sequence 2, Appl1  |
| 45 | 292.8 | 85.6 | 9511 | 1 | US-07-925-695-6   | Sequence 6, Appl1  |

## ALIGNMENTS

RESULT 1

Sequence 48, Application US/09014416

GENERAL INFORMATION:

APPLICANT: Bukh, Jens

APPLICANT: Purcell, Robert H.

1. TITLE OF INVENTION: USES THEREOF

CURRENT APPLICATION NUMBER: US/09/014,416  
CURRENT FILING DATE: 1009-01-27

EARLIER APPLICATION NUMBER: US 60/053,062  
EARLIER FILING DATE: 1997-07-18

```

; NUMBER OF SEQ ID NOS: 65
SOEIN/APE: Data to Ver 3.1

```

SEQ ID NO 48

ORGANISM: Hepatitis C virus

US-09-014-416-48

|                       |              |                    |               |             |
|-----------------------|--------------|--------------------|---------------|-------------|
| Query Match           | 95.1%        | Score 325.2;       | DB 3;         | Length 341; |
| Best Local Similarity | 98.8%        | Pred. No. 4.6e-99; |               |             |
| Matches 338;          | Conservative | 0;                 | Mismatches 3; | Indels 1;   |
|                       |              |                    | Gaps          | 1;          |

[illegible]

Db 300 ggtgcttgcgagtcgcccgagagcttgcgtagaccgtgaccc 341

## RESULT 2

US-08-432-693-1  
Sequence 1, Application US/08432693

Patent No. 5861267

GENERAL INFORMATION:

APPLICANT: Su, Michael

TITLE OF INVENTION: METHODS AND HOST CELLS FOR ASSAYING

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESS: Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/432,693

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Haley Jr, James F

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: VPI/95-01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-596-9090

TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9401 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 3420..5312

OTHER INFORMATION: /product= "NS3 protease"

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 5313..5474

OTHER INFORMATION: /product= "NS4A"

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 5475..5552

OTHER INFORMATION: /product= "truncated NS4B"

US-08-432-693-1

Query Match 95.1%; Score 325.2; DB 2; Length 9401;  
Best Local Similarity 98.8%; Pred. No. 1.8e-98;  
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Db 1 gccagcccccctgtaggggagacactccacatagatactccctgtgaggaactactg 60  
1 gccagcccccctgtaggggagacactccacatagatactccctgtgaggaactactg 60  
QY 1 tcttcacgcaaaagcgctcgtacatggttagatagatggtgctgagcctccagcc 120  
1 tcttcacgcaaaagcgctcgtacatggttagatagatggtgctgagcctccagcc 120  
Db 61 TCTTCACGCAAAAGCGCTGAGCCATGGCGTTAGTATAGTGTCTGTCGAGCCTCCAGGAC 120  
61 TCTTCACGCAAAAGCGCTGAGCCATGGCGTTAGTATAGTGTCTGTCGAGCCTCCAGGAC 120  
QY 121 ccccccctccgggagagccatagtggtctgcggaacggtgtagtacacggaaattgcca 180  
121 ccccccctccgggagagccatagtggtctgcggaacggtgtagtacacggaaattgcca 180

Db 131 ccccccctccgggagagccatagtggtctgcggaacggtgtagtacacggaaattgcca 180  
131 ccccccctccgggagagccatagtggtctgcggaacggtgtagtacacggaaattgcca 180

QY 131 gacgacccggtccttcttgatcaatcccgctcaatgctgagatggtggcggtccccc 240  
131 gacgacccggtccttcttgatcaatcccgctcaatgctgagatggtggcggtccccc 240

Db 131 GACGACCCGgtccttcttgatcaatcccgctcaatgctgagatggtggcggtccccc 239  
131 GACGACCCGgtccttcttgatcaatcccgctcaatgctgagatggtggcggtccccc 239

QY 241 cgcgagactgtagccagtagtggtggtcgcgaaagccttgtagtactgctgataag 300  
241 cgcgagactgtagccagtagtggtggtcgcgaaagccttgtagtactgctgataag 300

Db 240 CCGAAGACTGTAGCCGAGACTAGTGTGGGTCCGGAAGCCCTGTGTACTCCTCATAG 299  
240 CCGAAGACTGTAGCCGAGACTAGTGTGGGTCCGGAAGCCCTGTGTACTCCTCATAG 299

QY 301 ggtgcttgcgagtcgcccgagagcttgcgtagaccgtgaccc 342  
301 ggtgcttgcgagtcgcccgagagcttgcgtagaccgtgaccc 342

Db 300 GGTGCTTGCgagtcgcccgagagcttgcgtagaccgtgaccc 341  
300 GGTGCTTGCgagtcgcccgagagcttgcgtagaccgtgaccc 341

## RESULT 3

US-08-811-566-19

Sequence 19, Application US/08811566

Patent No. 6127116

GENERAL INFORMATION:

APPLICANT: Rice, Charles et al.

TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESS: David A. Jackson, Esq.

STREET: 411 Hackensack Ave, Continental Plaza, 4th

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/811,566

FILING DATE: 03-MAR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1113-1-006

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 9416 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

US-08-811-566-19

Query Match 95.1%; Score 325.2; DB 3; Length 9416;  
Best Local Similarity 98.8%; Pred. No. 1.8e-98;  
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Db 1 gccagcccccctgtaggggagacactccacatagatactccctgtgaggaactactg 60  
1 gccagcccccctgtaggggagacactccacatagatactccctgtgaggaactactg 60  
QY 61 tcttcacgcaaaagcgctcgtacatggttagatagatggtgctgagcctccagcc 120  
61 tcttcacgcaaaagcgctcgtacatggttagatagatggtgctgagcctccagcc 120  
Db 61 TCTTCACGCAAAAGCGCTGAGCCATGGCGTTAGTATAGTGTCTGTCGAGCCTCCAGGAC 120  
61 TCTTCACGCAAAAGCGCTGAGCCATGGCGTTAGTATAGTGTCTGTCGAGCCTCCAGGAC 120  
QY 121 ccccccctccgggagagccatagtggtctgcggaacggtgtagtacacggaaattgcca 180  
121 ccccccctccgggagagccatagtggtctgcggaacggtgtagtacacggaaattgcca 180

Db 121 CCCCCCTCCGGGAGAGCATAGTGTGCGGAACCGGTGATACCGGAATTGCCAG 180  
181 GACGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 239  
QY 241 CCGGAGACTGTCTAGCGGAGTAGTGTGGTCTCGGAAAGGCTGTGTACTGCTGATAG 300  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 240 CCGGAGACTGTCTAGCGGAGTAGTGTGGTCTCGGAAAGGCTGTGTACTGCTGATAG 299  
QY 301 GTGTCTGTGAGTGTGCGGAGGAGTGTCTGTAGACCGTGCACC 342  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 300 GGTGCTGTGAGTGTGCGGAGGAGTGTCTGTAGACCGTGCACC 341

RESULT 4  
US-09-014-416-4  
Sequence 4, Application US/09014416  
Patent No. 6153421  
GENERAL INFORMATION:  
APPLICANT: Yanagil, Masayuki  
APPLICANT: Bukh, Jens  
APPLICANT: Emerson, Susanne U.  
APPLICANT: Purcell, Robert H.  
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND  
FILE REFERENCE: 20264276  
CURRENT APPLICATION NUMBER: US/09/014,416  
CURRENT FILING DATE: 1998-01-27  
EARLIER APPLICATION NUMBER: US 60/053,062  
EARLIER FILING DATE: 1997-07-18  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 9595  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
US-09-014-416-4

Query Match 95.1%; Score 325.2; DB 3; Length 9595;  
Best Local Similarity 98.8%; Pred. No. 1.8e-98;  
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 1 GCCAGCCCCGTGATGGGCGGACACCCACCATAGATCCCTCCCTGTAGGAACTACTG 60  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 1 GCCAGCCCCGTGATGGGCGGACACCCACCATAGATCCCTCCCTGTAGGAACTACTG 60  
QY 61 TCTCAAGAGAAAGCGTCTAGCATGCGTGTAGTAGTGTCTGTGCGAGCTCCAGGCC 120  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 61 TCTCAAGAGAAAGCGTCTAGCATGCGTGTAGTAGTGTCTGTGCGAGCTCCAGGCC 120  
QY 121 CCCCCCTCCGGGAGAGCATAGTGTCTCGGAAACCGGTGATACCGGAATTGCCAG 180  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 121 CCCCCCTCCGGGAGAGCATAGTGTCTCGGAAACCGGTGATACCGGAATTGCCAG 180  
QY 181 GACGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 239  
Db 241 CCGGAGACTGTCTAGCGGAGTAGTGTGGTCTCGGAAAGGCTGTGTACTGCTGATAG 300  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 240 CCGGAGACTGTCTAGCGGAGTAGTGTGGTCTCGGAAAGGCTGTGTACTGCTGATAG 299  
QY 301 GTGTCTGTGAGTGTGCGGAGGAGTGTCTGTAGACCGTGCACC 342  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 300 GGTGCTGTGAGTGTGCGGAGGAGTGTCTGTAGACCGTGCACC 341

RESULT 5  
US-09-014-416-6  
Sequence 6, Application US/09014416

Patent No. 6153421  
GENERAL INFORMATION:  
APPLICANT: Yanagil, Masayuki  
APPLICANT: Bukh, Jens  
APPLICANT: Emerson, Susanne U.  
APPLICANT: Purcell, Robert H.  
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND  
FILE REFERENCE: 20264276  
CURRENT APPLICATION NUMBER: US/09/014,416  
CURRENT FILING DATE: 1998-01-27  
EARLIER APPLICATION NUMBER: US 60/053,062  
EARLIER FILING DATE: 1997-07-18  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 6  
LENGTH: 9599  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
US-09-014-416-6  
Query Match 95.1%; Score 325.2; DB 3; Length 9599;  
Best Local Similarity 98.8%; Pred. No. 1.8e-98;  
Matches 338; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 1 GCCAGCCCCGTGATGGGCGGACACCCACCATAGATCCCTCCCTGTAGGAACTACTG 60  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 1 GCCAGCCCCGTGATGGGCGGACACCCACCATAGATCCCTCCCTGTAGGAACTACTG 60  
QY 61 TCTCAAGAGAAAGCGTCTAGCATGCGTGTAGTAGTGTCTGTGCGAGCTCCAGGCC 120  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 61 TCTCAAGAGAAAGCGTCTAGCATGCGTGTAGTAGTGTCTGTGCGAGCTCCAGGCC 120  
QY 121 CCCCCCTCCGGGAGAGCATAGTGTCTCGGAAACCGGTGATACCGGAATTGCCAG 180  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 121 CCCCCCTCCGGGAGAGCATAGTGTCTCGGAAACCGGTGATACCGGAATTGCCAG 180  
QY 181 GACGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 239  
Db 241 CCGGAGACTGTCTAGCGGAGTAGTGTGGTCTCGGAAAGGCTGTGTACTGCTGATAG 300  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 240 CCGGAGACTGTCTAGCGGAGTAGTGTGGTCTCGGAAAGGCTGTGTACTGCTGATAG 299  
QY 301 GTGTCTGTGAGTGTGCGGAGGAGTGTCTGTAGACCGTGCACC 342  
181 GAGGACGGGTCCTCTCTGATCAATCCGCTCAATGCTGAGATTGGGCGTCCCC 240  
Db 300 GGTGCTGTGAGTGTGCGGAGGAGTGTCTGTAGACCGTGCACC 341

RESULT 6  
US-09-014-416-6  
Sequence 1, Application US/08440209  
Patent No. 5922857  
GENERAL INFORMATION:  
APPLICANT: Han, Jang H.  
APPLICANT: Spaete, Richard R.  
TITLE OF INVENTION: Methods and Compositions for Controlling  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESS: Wolf, Greenfield, and Sacks P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,209  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/477,895  
FILING DATE: 06-JUN-1995  
APPLICATION NUMBER: US 08/128,583  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Janluk, Anthony J  
REGISTRATION NUMBER: 29,809  
REFERENCE/DOCKET NUMBER: C0772/7004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: YES  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: ns5hcv1  
US-08-440-209-1

Query Match 94.6%; Score 323.6; DB 2; Length 341;  
Best Local Similarity 98.5%; Pred. No. 1.6e-98;  
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 1 gccagccccctatgaggcgagacatccacatagatcaccctctgaggaactactg 60  
DB 1 gccacccccctgagggcgagacatccacatagatcaccctctgaggaactactg 60  
QY 61 tcttacgcaaaagcgcctagccatgagctagatagatgagctgagcctccagcc 120  
DB 61 tcttacgcaaaagcgcctagccatgagctagatagatgagctgagcctccagcc 120  
QY 121 cccccctccggagagacatagctgctcgagaaacggtagacacgggaattgcca 180  
DB 121 cccccctccggagagacatagctgctcgagaaacggtagacacgggaattgcca 180  
QY 181 gacgacgggtctcttcttgatcaatcccgctcaatgctgaggaattggcgctgcc 240  
DB 181 gacgacgggtctcttcttgatcaatcccgctcaatgctgaggaattggcgctgcc 240  
QY 241 cgcgagactgtagccgagtagtgtgtggtcgagaaagccttgtagctgctgata 300  
DB 241 cgcgagactgtagccgagtagtgtgtggtcgagaaagccttgtagctgctgata 300  
QY 301 ggtgctgcagtgcccgagaggtctcgtagaccgtgcacc 342  
DB 301 ggtgctgcagtgcccgagaggtctcgtagaccgtgcacc 342

RESULT 7  
US-08-854-531-4  
Sequence 4, Application US/08854531  
Patent No. 6025341  
GENERAL INFORMATION:  
APPLICANT: Wands, Jack  
APPLICANT: Tokushige, Katsutoshi  
APPLICANT: Wakita, Takaji  
TITLE OF INVENTION: CHIMERIC HEPATITIS B/HEPATITIS C VIRUS VACCINE  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6025341.ris  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia

STATE: PA.  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,531  
FILING DATE: Herewith  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: APOL-0214  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-854-531-4

Query Match 94.6%; Score 323.6; DB 3; Length 341;  
Best Local Similarity 98.5%; Pred. No. 1.6e-98;  
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 1 gccagccccctatgaggcgagacatccacatagatcaccctctgaggaactactg 60  
DB 1 gccacccccctgagggcgagacatccacatagatcaccctctgaggaactactg 60  
QY 61 tcttacgcaaaagcgcctagccatgagctagatagatgagctgagcctccagcc 120  
DB 61 tcttacgcaaaagcgcctagccatgagctagatagatgagctgagcctccagcc 120  
QY 121 cccccctccggagagacatagctgctcgagaaacggtagacacgggaattgcca 180  
DB 121 cccccctccggagagacatagctgctcgagaaacggtagacacgggaattgcca 180  
QY 181 gacgacgggtctcttcttgatcaatcccgctcaatgctgaggaattggcgctgcc 240  
DB 181 gacgacgggtctcttcttgatcaatcccgctcaatgctgaggaattggcgctgcc 240  
QY 241 cgcgagactgtagccgagtagtgtgtggtcgagaaagccttgtagctgctgata 300  
DB 241 cgcgagactgtagccgagtagtgtgtggtcgagaaagccttgtagctgctgata 300  
QY 301 ggtgctgcagtgcccgagaggtctcgtagaccgtgcacc 342  
DB 301 ggtgctgcagtgcccgagaggtctcgtagaccgtgcacc 342

RESULT 8  
US-08-439-996-1  
Sequence 1, Application US/08439996  
Patent No. 6057093  
GENERAL INFORMATION:  
APPLICANT: Han, Jang H  
APPLICANT: Spaete, Richard R  
TITLE OF INVENTION: Methods and Compositions for Controlling  
TITLE OF INVENTION: Translation of HCV proteins  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield, and Sacks P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA

```

RESULT 9
US-09-014-416-47
; Sequence 47, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagii, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.

```

US-08-869-380-4  
Sequence 4, Application US/0869380  
Patent No. 6235888  
GENERAL INFORMATION:  
APPLICANT: Pachuk, Catherine J.  
APPLICANT: Mandis, Jack  
APPLICANT: Makita, Takaji  
APPLICANT: Zurawski, Jr., Vincent R.  
APPLICANT: Coney, Leslie R.  
TITLE OF INVENTION: Hepatitis C Virus Vaccine  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &  
ADDRESSEE: No. 6235888ris  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: PA.  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/869,380





STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 555X  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,700B  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/240,382  
FILING DATE: 10 May 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 00786/279001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ. ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-474-700B-39

Query Match 94.6%; Score 323.6; DB 3; Length 342;  
Best Local Similarity 98.5%; Pred. No. 1.6e-98;  
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 1 gccagccccctgatatggggcgacactccaccatagatcactccctgtgaggaactactg 60  
DB 342 gccagccccctgatatggggcgacactccaccatagatcactccctgtgaggaactactg 283  
QY 61 tcttcaagcagaagcgctcctagcctatgctatgagtgtctgcaagctccagcc 120  
DB 282 tcttcaagcagaagcgctcctagcctatgctatgagtgtctgcaagctccagcc 223  
QY 121 ccccccccgagagagcacaatgctcgcgagacggctgagatcccggaattgcag 180  
DB 222 ccccccccgagagagcacaatgctcgcgagacggctgagatcccggaattgcag 163  
QY 181 gacgacgggctccttcttgatcaatccgcctcaatgctgagattggcgctgc 240  
DB 162 gacgacgggctccttcttgatcaatccgcctcaatgctgagattggcgctgc 104  
QY 241 cgcgagactctagcagatgctgtgctcgcgaaagcctgtgtactgctgtag 300  
DB 103 cgcgagactctagcagatgctgtgctcgcgaaagcctgtgtactgctgtag 44  
QY 301 ggtgcttcgagtgcccccgagaggtctcgtagaccgtgcac 342  
DB 43 ggtgcttcgagtgcccccgagaggtctcgtagaccgtgcac 2

RESULT 13  
US-07-863-622-1  
Sequence 1, Application US/07863622  
PATENT NO. 5858650  
GENERAL INFORMATION:  
APPLICANT: CELEBUSKI, JOSEPH E.  
TITLE OF INVENTION: METHODS FOR INACTIVATING NUCLEOTIDE  
NUMBER OF INVENTION: SEQUENCES AND METAL CHELATES FOR USE THEREIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:

ADDRESSEE: ABBOTT LABORATORIES  
STREET: D377/APED, ONE ABBOTT PARK ROAD.  
CITY: ABBOTT PARK  
STATE: ILLINOIS  
COUNTRY: LAKE  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/863,622  
FILING DATE: 19920403  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WONG, WEAN KHING  
REGISTRATION NUMBER: 33561  
REFERENCE/DOCKET NUMBER: 5159.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-938-3517  
TELEFAX: 708-937-9556  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 350 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-07-863-622-1

Query Match 94.6%; Score 323.6; DB 2; Length 350;  
Best Local Similarity 98.5%; Pred. No. 1.6e-98;  
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 1 gccagccccctgatatggggcgacactccaccatagatcactccctgtgaggaactactg 60  
DB 342 gccagccccctgatatggggcgacactccaccatagatcactccctgtgaggaactactg 60  
QY 61 tcttcaagcagaagcgctcctagcctatgctatgagtgtctgcaagctccagcc 120  
DB 61 tcttcaagcagaagcgctcctagcctatgctatgagtgtctgcaagctccagcc 120  
QY 121 ccccccccgagagagcacaatgctcgcgagacggctgagatcccggaattgcag 180  
DB 121 ccccccccgagagagcacaatgctcgcgagacggctgagatcccggaattgcag 180  
QY 181 gacgacgggctccttcttgatcaatccgcctcaatgctgagattggcgctgc 240  
DB 181 gacgacgggctccttcttgatcaatccgcctcaatgctgagattggcgctgc 239  
QY 241 cgcgagactctagcagatgctgtgctcgcgaaagcctgtgtactgctgtag 300  
DB 240 cgcgagactctagcagatgctgtgctcgcgaaagcctgtgtactgctgtag 299  
QY 301 ggtgcttcgagtgcccccgagaggtctcgtagaccgtgcac 342  
DB 300 ggtgcttcgagtgcccccgagaggtctcgtagaccgtgcac 341

RESULT 14  
PCT-US93-03266-1  
Sequence 1, Application PC/TUS9303266  
GENERAL INFORMATION:  
APPLICANT: CELEBUSKI, JOSEPH E.  
TITLE OF INVENTION: METHODS FOR INACTIVATING NUCLEOTIDE  
NUMBER OF INVENTION: SEQUENCES AND METAL CHELATES FOR USE THEREIN  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ABBOTT LABORATORIES  
STREET: D377/APED, ONE ABBOTT PARK ROAD  
CITY: ABBOTT PARK

```

STATE: ILLINOIS
COUNTRY: LAKE
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03266
FILING DATE: 19930405
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MONG MEAN KHING
REGISTRATION NUMBER: 33561
REFERENCE/DOCKET NUMBER: 5159.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-938-3517
TELEFAX: 708-938-2623
SEQUENCE CHARACTERISTICS:
LENGTH: 350 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-03266-1

```

```

Query Match 94.6% Score 323.6; DB 5; Length 350;
Best Local Similarity 98.5% Pred. No. 1.6e-98;
Matches 337; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```

```

QY 1 gccagccctctatgaggcgagacatccacatagatcctccctgtgaggaactactg 60
 |||
DB 1 gccagccctctatgaggcgagacatccacatagatcctccctgtgaggaactactg 60
 |||
QY 61 tcttcacgagaagaagcgtctagccatgagcgttagatgtgtgagcagccctccagcc 120
 |||
DB 61 tcttcacgagaagaagcgtctagccatgagcgttagatgtgtgagcagccctccagcc 120
 |||
QY 121 ccccccctccggagagacatagtgctgcggaaccggtgagtaaccggaattgcag 180
 |||
DB 121 ccccccctccggagagacatagtgctgcggaaccggtgagtaaccggaattgcag 180
 |||
QY 181 gagcagcgggtccttcttgatcaatcccgctcaatgccttgagagattggcggtcc 240
 |||
DB 181 gagcagcgggtccttcttgatcaatcccgctcaatgccttgagagattggcggtcc 240
 |||
QY 241 cgcgagacgtctagccgagtagtgttgggtcgcaaaagccttggtgtactcctgatat 300
 |||
DB 241 cgcgagacgtctagccgagtagtgttgggtcgcaaaagccttggtgtactcctgatat 300
 |||
QY 301 ggtgcttcgagtgcccgaggagctcgtagaccgtgcac 342
 |||
DB 301 ggtgcttcgagtgcccgaggagctcgtagaccgtgcac 342
 |||
QY 300 ggtgcttcgagtgcccgaggagctcgttagaccgtgcac 341
 |||
DB 300 ggtgcttcgagtgcccgaggagctcgttagaccgtgcac 341
 |||

```

```

RESULT 15
US-08-988-321B-37
Sequence 37, Application US/08988321B
Patent No. 6174688
GENERAL INFORMATION:
APPLICANT: Kevin P. Anderson et al.
TITLE OF INVENTION: Compositions And Methods For Treatment Of Hepatitis C V
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESS: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Milton
STATE: NJ
COUNTRY: USA
ZIP: 08053

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORDPERFECT 6.1 FOR WINDOWS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/988,321B
FILING DATE: December 10, 1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/650,093
FILING DATE: May 17, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/452,841
FILING DATE: May 30, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/397,220
FILING DATE: March 9, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/945,289
FILING DATE: September 10, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0245
TELECOMMUNICATION INFORMATION:
TELEPHONE: (856) 810-1515
TELEFAX: (856) 810-1454
SEQUENCE CHARACTERISTICS:
LENGTH: 686
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: No
US-08-988-321B-37

```

```

Query Match 94.6% Score 323.6; DB 4; Length 686;
Best Local Similarity 78.9% Pred. No. 2.1e-98;
Matches 270; Conservative 67; Mismatches 4; Indels 1; Gaps 1;

```

```

QY 1 gccagccctctatgaggcgagacatccacatagatcctccctgtgaggaactactg 60
 |||
DB 1 gccagccctctatgaggcgagacatccacatagatcctccctgtgaggaactactg 60
 |||
QY 61 tcttcacgagaagaagcgtctagccatgagcgttagatgtgtgagcagccctccagcc 120
 |||
DB 61 tcttcacgagaagaagcgtctagccatgagcgttagatgtgtgagcagccctccagcc 120
 |||
QY 121 ccccccctccggagagacatagtgctgcggaaccggtgagtaaccggaattgcag 180
 |||
DB 121 ccccccctccggagagacatagtgctgcggaaccggtgagtaaccggaattgcag 180
 |||
QY 181 gagcagcgggtccttcttgatcaatcccgctcaatgccttgagagattggcggtcc 240
 |||
DB 181 gagcagcgggtccttcttgatcaatcccgctcaatgccttgagagattggcggtcc 240
 |||
QY 241 cgcgagacgtctagccgagtagtgttgggtcgcaaaagccttggtgtactcctgatat 300
 |||
DB 241 cgcgagacgtctagccgagtagtgttgggtcgcaaaagccttggtgtactcctgatat 300
 |||
QY 301 ggtgcttcgagtgcccgaggagctcgtagaccgtgcac 342
 |||
DB 301 ggtgcttcgagtgcccgaggagctcgttagaccgtgcac 341
 |||
QY 300 ggtgcttcgagtgcccgaggagctcgttagaccgtgcac 341
 |||
DB 300 ggtgcttcgagtgcccgaggagctcgttagaccgtgcac 341
 |||

```

```

Search Completed: March 1, 2002, 15:26:27
Job time: 36645 sec

```

---

THIS PAGE BLANK (USPTO)

**THIS PAGE BLANK (USPTO)**